

OM of: US-08-900-220C-17_COPY_20_198 to: Issued_Patents_NA.*

out_format : pfs

Date: Apr 22, 2001 12:05 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlp
-O=/cgn2_1/USPTO_spool/US08900220/runat.19042001.104047.21686/app.query.fasta.1.253
-DB=Issued_Patents_NA -OEMT=fasta -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOEPEL=0.000 -LOEPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELDP=6.000 -DELTEXT=7.000 -START=1 -MATRIX=blsum62
-TRAX=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs
-NORMEXT=MINLEN=0 -MAXLEN=2000000000
-USER=US08900220 -CGN1_1_55 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREDS=1

Search information block:

Query: US-08-900-220C-17_COPY_20_198
Query length: 179
Database: Issued_Patents_NA.*
Database sequences: 302621
Database length: 87301344
Search time (sec): 80.530000

score_list:

Sequence	Strd Orig	ZScore	EScore	len	Documentation	..	2.8e-103	1190
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-1				741.00	1575.99	3.4e-80	1277	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				741.00	1575.99	3.4e-80	1277	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				740.00	1573.52	4.6e-80	1313	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-356-060A-4				740.00	1573.52	4.6e-80	1313	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				740.00	1573.52	4.6e-80	1313	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				738.00	1568.28	9.1e-80	1425	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				738.00	1568.28	9.1e-80	1425	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-10				734.00	1558.50	3.2e-79	1576	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-15				732.00	1551.17	5.5e-79	1576	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				729.50	1551.06	8.3e-79	1281	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-9				723.00	1537.21	4.9e-78	1256	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-5				723.00	1537.21	4.9e-78	1256	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-5				723.00	1537.21	4.9e-78	1256	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:PCR-US95-02315-1				718.00	1522.93	3.0e-77	1715	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-356-060A-3				569.00	1205.73	1.4e-59	1056	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-7				462.00	975.38	9.6e-47	939	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-7				462.00	975.38	9.6e-47	939	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-2				238.00	511.25	6.8e-21	144	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-567-357A-2				238.00	511.25	6.8e-21	144	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:PCR-US95-15463-2				238.00	511.25	6.8e-21	144	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:PCR-US95-15923-2				226.00	485.27	1.9e-19	144	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				226.00	485.27	1.9e-19	144	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-567-357A-1				226.00	485.27	1.9e-19	144	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:PCR-US95-15463-1				226.00	485.27	1.9e-19	144	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:PCR-US95-15923-1				82.00	133.81	7.22	5141	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				82.00	133.81	7.22	5141	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-764-233A-1				81.00	128.63	3.89	4937	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-09-332-468A-1				80.00	130.02	11.73	2742	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-09-332-468A-1				80.00	130.02	11.73	2742	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-795-868-15				79.00	114.83	6.34	2614	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-928-934-28				77.50	117.94	5.52	8931	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-648-657-14				76.00	133.20	7.80	1686	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-648-657-14				76.00	133.20	7.80	1686	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:PCR-US95-14418-3				76.00	132.52	8.52	1794	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:PCR-US95-15327-3				76.00	132.52	8.52	1794	1

/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-073-384C-3 + 76.00 128.81 13.71 2504
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-254-359A-3 + 76.00 128.81 13.71 2504
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-483-043-3 + 76.00 128.81 13.71 2504
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-481-238-3 + 76.00 128.81 13.71 2504
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-471-066B-3 + 76.00 128.81 13.71 2504

seq_name: /cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-3

seq_documentation block:

Sequence 3, Application US/08176427B
Patent No. 5785543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related thereto
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 1190 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1191
US-08-176-427B-3

alignment_scores:

Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

Alignment block:

US-08-900-220C-17_COPY_20_198 x US-08-176-427B-3 ..

Align seg 1/1 to: US-08-176-427B-3 from: 1 to: 1190

1 AAGinscGcGlyProGlyArgGlyProValGlyArgArgGlyTyrAl 17
|||||
58 GCCCAGAGTGGCGGGCGGCGGAGGACCGGTTGGCGGCGGCTATAT 107
17 aatgtyGcGlnValProLeuLeuTyrTrpGlnPheValProGlyValP 34
:|||||
108 GCGCAGACACTTGTCCCTCTGCTATACACAGCATTGGCCGACAGATGC 157
34 rogluarGthrleuGlalaserGlyProAlaGluGlyArgValAlaArg 50

```
|||||
158 CCGAGCGGACCTGGGCGGAGTGGCCAGCGAGGAGGGATACAGG 207
51 GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
208 GGGTCGGAGGCTTCGGGACCTCGTACCCACTACACCCGACATAT 257
67 ePhelysAspGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAGGATGAGGAGACAGCGCGGACGCCCTGTATGACAGCGCTT 307
84 yslYsgIuArgValAsnAlaLeuAlaIleAlaValMetAsnMetPro 100
|||||
308 GCNAAGGCGGGTGAACGCTTAACCATCGCGGATGATGATGGCC 357
101 GlyValArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHSAI 117
|||||
358 GGAGTACGCTAGTGTGACTGAAGGCTGGGAGAGGAGCCGACACG 407
117 agInAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThSera 134
|||||
408 ACAGATTCACTCCACTACAGAGCGCGCTTGACATCACACGCTTG 457
134 spArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGTGACCGTAATAGATGATGTTGTGGCGGCGCTAGCTGTGAAG 507
151 GlyPheAspTrrValTyrTyrGluSerArgAsnHisValHisValSera 167
|||||
508 GGATTGACAGGGGTCTACTAGAGTCCCGCACACACATCCACGATG 557
167 llySAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
558 CAAAGCTGATTACTCACTGCGGCGCGAGCGCGAGGC 594

seq_name: /cgn2_6/plodata/2/lna/5B_COMB.seq:us-08-356-060A-2

seq_documentation_block:
; Sequence 2, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1190 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1191
US-08-356-060A-2

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:
US-08-900-220c-17_COPY_20_198 x US-08-356-060A-2 ..
Align seq 1/1 to: US-08-356-060A-2 from: 1 to: 1190

1 AlAGlnSerGysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
|||||
58 GCCCAGAGCTGCGGGCGCGGCGAGAGACCGGTTGGCGGCGGCTTATGT 107
17 aArgLysGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
|||||
108 GCCGACACCACTTGTGCTCTGCTATCAAGCAGATTGTGCCAGTATGC 157
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CCGAGCGGACCTGCGGCGCGAGTGGCCAGCGGAGGAGGATGAACAAG 207
51 GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
208 GGGTCGGAGGCGCTTCGGGACCTGTAACCCACTACAAACCCGACATAT 257
67 ePhelysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAGAGATGAGGAGAGAACAGCGCGGACGCGCTGTATGACAGAGGCT 307
84 yslYsgIuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGAGACCGGTGAACGCTCTAGCCATCCGCGGTGATGACATGTGCC 357
101 GlyValArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHSAI 117
|||||
358 GGAGTACGCTTACGTGTGACTGAAGGCTGGGAGGAGGCGCACACGCG 407
117 agInAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThSera 134
|||||
408 ACAGATTCACTCCACTACAGAGCGCGCTTGAGCATCACACGCTTG 457
134 spArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGTGACCGTAATAGATGATGTTGTGGCGGCGCTAGCTGTGAAGCC 507
151 GlyPheAspTrrValTyrTyrGluSerArgAsnHisValHisValSera 167
|||||
508 GGATTGACAGGGGTCTACTAGAGTCCCGCACACACATCCACGATG 557
167 llySAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
558 CAAAGCTGATTACTCACTGCGGCGCGAGCGCGAGGC 594

seq_name: /cgn2_6/plodata/2/lna/6B_COMB.seq:us-08-460-900C-2

seq_documentation_block:
; Sequence 2, Application US/08460900C
; Patent No. 6165747
; GENERAL INFORMATION:
```

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
APPLICANT: Buncroft, David A.
APPLICANT: Marti-Gorostiza, Elisa
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,900C
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 4-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1190 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1188
US-08-460-900C-2

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:
US-08-900-220C-17_COPY_20_198 x US-08-460-900C-2 ..
Align seg 1/1 to: US-08-460-900C-2 from: 1 to: 1190

1 AAgInSerCysGlyProGlyArgValGlyArgArgTyrAl 17
|||||
58 GCGGAGAGCTGGGGCGGCGGAGAGCGGTGGCGCGCGCTATGT 107
17 ArgGlySerLeuValProLeuLeuTyrGlyGlnPheValProGlyValP 34
|||||
108 GCGCAGAGCAACTGTGCTCTGCTATACAGACGTTGTGCGCCACTATGC 157
34 roGluArgThrLeuGluAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CCGAGCGGACCTGGGCGCGGAGTGGCCAGCGGAGGAGGAGGTAAACAAAG 207

51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle1 67
|||||
208 GCGTCGAGCGCTTCGCGGAGCTGCTACCCACTACACACCGGACATAT 257
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAAGCATGAGAGAACAGCGCGGAGCGGCGCTGATGACAGAGGTT 307
84 ySLysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGAGCGGCTGAACGCTTACGCAATCGCGGTATGAACTGTGGCC 357
101 GlyValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAl 117
|||||
358 GGAGTACCGCTACGTGCTGACTGAAGCTGGGAGCAGAGCGGACACAGC 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSerA 134
|||||
408 ACAGGATTCACCTCCACTACGAAGCGGCTGTGACATCACACAGCTTG 457
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAl 150
|||||
458 ACCGTACCGCTATATAGTATGTTGTGTGGCGGCTGACTGTGGAAGCC 507
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerA 167
|||||
508 GGATTCGACTGGCTGCTACAGCAGTCCGCGACACACATCCACTATCGT 557
167 lLysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||
558 CAAAGCTGATTAACCTGCGGCGGCGGAGCCGAGGCG 594

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:us-08-176-427B-1

seq_documentation_block:
Sequence 1, Application US/08176427B
Patent No. 5789543
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1277 base pairs
TYPE: nucleic acid
STRANDEDNESS: both

TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1277
US-08-176-427B-1

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.411 Gaps: 2
Percent Similarity: 94.915 Percent Identity: 74.011

alignment_block:

US-08-900-220c-17_COPY_20_198 x US-08-176-427B-1 ..

Align seg 1/1 to: US-08-176-427B-1 from: 1 to: 1277

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgArgGly 19
  :::::::::::::::::::::
76 ACTGTGTGACCAAGGAGGAGG...ATTGGAAAAAGAGGACCCCAAAA 122
  19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGly 36
  | :::::::::::::::::::::
123 G...CTGACCCCGTTAGCCTATAGCAGTTTATCCCAATGTGGCAGAGA 169
  36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
  :::::::::::::::::::::
170 AGACCCCTAGGAGGCGCAGTGAAGATGTGAAGGAGGATCACAGAACTCC 219
  53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
  :::::::::::::::::::::
220 GAGGATTTAAAGAACTAACCCCAATTACAACTGACATTTATTTTAA 269
  69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLys 86
  :::::::::::::::::::::
270 GGATGAGAGAGACACGCGACCTGCACATGATGACTACCGCGCCAGG 319
  86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  :::::::::::::::::::::
320 ACAAGCTGAATGCCCTGGCGATCTGGGTGATGAACAGTGCGCGGGGTG 369
  103 ArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHisAlaGlnAs 119
  :::::::::::::::::::::
370 AAGCTGGGGTGAAGGAGGCTGGGAGAGAGATGCCATCTCCGAGGA 419
  119 pSerLeuHisTyrGluGluArgAlaLeuAspIleThrThrSerAspArg 136
  :::::::::::::::::::::
420 ATGCTCTCATTAGSAGGCTGCGCCCTGCGACATCCACACGTCGATCGG 469
  136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
  :::::::::::::::::::::
470 ACCGACACAGATAGCGAATGCTGCGCCCTGCGAGGCGGCGCTTC 519
  153 AspThrValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
  :::::::::::::::::::::
520 GACTGGGTCTACTACGAGTCCAGAGCGCAATCTCAGTCTCCGTAAGC 569
  169 aAspAsnSerLeuAlaValArgAlaGlyLys 179
  :::::::::::::::::::::
570 AGAAACTCAGTGGCAGCAAAATCAGAGG 600
  seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-356-060A-1
seq_documentation_block:
; Sequence 1, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto

```

```

; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMT-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1275
; US-08-356-060A-1

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.411 Gaps: 2
Percent Similarity: 94.915 Percent Identity: 74.011

alignment_block:
US-08-900-220c-17_COPY_20_198 x US-08-356-060A-1 ..
Align seg 1/1 to: US-08-356-060A-1 from: 1 to: 1277

3 SerCysGlyProGlyArgGlyProValGlyArgArgArgGly 19
  :::::::::::::::::::::
76 ACTGTGTGACCAAGGAGGAGG...ATTGGAAAAAGAGGACCCCAAAA 122
  19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGly 36
  | :::::::::::::::::::::
123 G...CTGACCCCGTTAGCCTATAGCAGTTTATCCCAATGTGGCAGAGA 169
  36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
  :::::::::::::::::::::
170 AGACCCCTAGGAGGCGCAGTGAAGATGTGAAGGAGGATCACAGAACTCC 219
  53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
  :::::::::::::::::::::
220 GAGGATTTAAAGAACTAACCCCAATTACAACTGACATTTATTTTAA 269
  69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLys 86
  :::::::::::::::::::::
270 GGATGAGAGAGACACGCGACCTGCACATGATGACTACCGCGCCAGG 319
  86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  :::::::::::::::::::::
320 ACAAGCTGAATGCCCTGGCGATCTGGGTGATGAACAGTGCGCGGGGTG 369

```

```

103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AAGCTGCGGTGACCGAGCGCTGGGACGACGATGCGCATCATCCGAGAGA 419
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgA 136
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATCGCTGACACAGAGGCTCGCGCGGTGACATCCACACGTCGGATCGGG 469
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCGCGAGAACTACGGAATGCTGGCCCGCTCGCGGTGAGCGCGCTTC 519
153 AsPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GACGTGGCTCTACTACGACTCCAGCGCGCACATCCACTGCTCGCTCAAGC 569
169 aAspAsnSerLeuAlaValArgAlaGlyLys 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 AGAAAACTCAGTGGCGAGCAATCATCGAGGCG 600

seq_name: /cgn2_6/ptodata/2/lna/5a_COMB.seq:us-08-460-900c-1

seq_documentation_block:
; Sequence 1, Application US/08460900C
; Patent No. 6165747
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; APPLICANT: Bumcrot, David A.
; APPLICANT: Marti-Gorostiza, Elisa
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,900C
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 4-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1275
; US-08-460-900c-1

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.41 Gaps: 2
Percent Similarity: 94.915 Percent Identity: 74.011

alignment_block:
US-08-900-220c-17_copy_20_198 x US-08-460-900c-1 ..
Align seg 1/1 to: US-08-460-900c-1 from: 1 to: 1277

3 SerCysGlyProGlyArgGlyProValGlyArgGlyArgTyrAlaArgLys 19
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
76 ACTGTGTGGACGACGAGGAGGCG...ATTGGMAAAGAGAGACCCCAAAA 122
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGluA 36
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 G...CTGACCCCGTTAGCCTATACGACGTTATTCCTCAATGTGGCAGAGA 169
36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgLysSer 52
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 AGACCCCTAGGGGCGCAGTGAAGATATGAAGGAGACATCCAGAAACTCC 219
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLys 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 GAGAGATTAAAGAACTAACCCCAATTAACAACCTGACATTATTTTAA 269
69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 GGATTAAGACGACGCGGAGCTGACAGACTGATGACTAGCCCTGACAGG 319
86 LuArgValAsnAlaLeuAlaValAlaValMetAsnMetThrProGlyVal 102
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ACAAGCTGAATGCCCTGGCGGATCTCGGTGATGACACAGTGGCGGGGG 369
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AAGCTGCGGTGACCGAGGCTGGGACGAGATGGCCATCATCTCCGAGAGA 419
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgA 136
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATCGCTGACACAGAGGCTCGCGCGGTGACATCCACACGTCGGATCGGG 469
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCGCGAGAACTACGGAATGCTGGCCCGCTCGCGGTGAGCGCGCTTC 519
153 AsPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GACTGGGTCTACTACGAGTCCAGGCGCACATCCACTGCTCGCTCAAGC 569
169 aAspAsnSerLeuAlaValArgAlaGlyLys 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 AGAAAACTCAGTGGCGAGCAATCATCGAGGCG 600

seq_name: /cgn2_6/ptodata/2/lna/5a_COMB.seq:us-08-176-427b-7

seq_documentation_block:
; Sequence 7, Application US/08176427B
; Patent No. 5789543
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMI-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-176-427B-7

```

```

alignment_scores:
Quality: 740.00      Length: 177
Ratio: 4.431         Gaps: 2
Percent Similarity: 94.350   Percent Identity: 74.576

```

alignment_block:
US-08-900-220C-17_COPY_20_198 x US-08-176-427B-7 ..

Align seg 1/1 to: US-08-176-427B-7 from: 1 to: 1313

```

3 SerGcSLyPProGlyArGlyProValGlyArGArGArGtyrAlaArgly 19
  ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|
70 GCCTGTGGCGCCGCGCAGGGG...TTTGAAGAAGCGCGCACCCCAAAA 116
19 scGlnLeuValProLeuLeuTyrlysglnPheValProGlyValProGln 36
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
117 G...CTGACCCCTTACCTACAGCAGATTATTCCCAACGTACCCGAGA 163
36 rGThrLeuGlyAlaSerGlyProAlaGlnGlyArGValAlaArglySer 52
  ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|
164 AGACCCAGGGGCGCGCAGCATATGAGGAAGATCACAGAAATCC 213
53 GlnArGpHeArGAsPLeuValProAsnTyrAsnProAspIleIlePhe 69
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
214 GAACGATTTAAGAACTACCCCAATTACACCCGACATCATATTAA 263
69 sAspGlnGlnAsnSerGlyAlaAspArgLeuMetThrGlnArgCysArg 86
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
264 GGATGAGCAAAACACGGGACGACGCGGTGATGACTCAGAGGGCAAG 313
86 LuArGValAsnAlaLeuAlaIleAlaValMetAsnMetTyrProGlyVal 102
  ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|
314 ACAAGTTAAATGCTTGCCATCTCTGTGATGAACAGTGCCTGGAGTG 363
103 ArgLeuArGValThrGlnGlyTyrPAspGlnAspGlyHisHisAlaGln 119
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
364 AGCTGCGAGTGAACGAGGCTGGATGAGACGCGCATCTTACAGAGA 413

```

```

119 pSerLeuHisTyrGlnGlyArGAlaLeuAspIleThrThrSerAspArg 136
  ::::::::::::::::::::|::|::|::|::|::|::|::|::|::|::|
414 GTCTTAACACTATGAGGTCGACAGTACATCCACCATCGGCGG 463
136 sPArGAsnLySTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGlyPhe 152
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
464 ACCGACAGATGACGCGCATCTGCTCGCTGCTGCGTGGAAAGGTTTC 513
514 GACTGGGCTACTATGATGAATCCAAGCTCACATCCACTGTTCTGTGAAGC 563
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
  | | | | | | | | | | | | | | | | | | | | | | | | | | |
564 AGAAGACTCGTGGCGGCAAAATCCGCGGCGC 594

```

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-08-356-060A-4

seq_documentation_block:

Sequence 4, Application US/08356060A

Patent No. 5844079

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.

APPLICANT: Tablin, Clifford J.

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

NUMBER OF SEQUENCES: 47

CORRESPONDENCE ADDRESS:

ADDRESSER: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,060A

FILING DATE: 14-DEC-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/176,427

FILING DATE: 30-DEC-1993

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.

REGISTRATION NUMBER: 36,709

REFERENCE/DOCKET NUMBER: HMI-006CP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1313 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1314

US-08-356-060A-4

```

alignment_scores:
Quality: 740.00      Length: 177
Ratio: 4.431         Gaps: 2
Percent Similarity: 94.350   Percent Identity: 74.576

```

alignment_block:

US-08-900-220C-17_COPY_20_198 x US-08-356-060A-4 ..

Align seg 1/1 to: US-08-356-060A-4 from: 1 to: 1313

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaArgLy 19
  ::::::::::::::::::::
70 GCCGTGGGCCCCGAGGAGG...TTGGAAAGAGGCGGACCCCAAAA 116
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGln 36
  ::::::::::::::::::::
117 G...CTGACCCCTTACGCTACAGCAGATTATTTCCACAGCTACCGAGA 163
36 rGTrLeuGlyAlaSerGlyProAlaGlyArgValAlaArgLySer 52
  ::::::::::::::::::::
164 AGACCCAGAGGCGGCGGAGATGAGAGGAGATCAAGAAAGATCC 213
53 GluArgPheArgAspLeuValProAsnTyrTrsProAspIleIlePhe 69
  ::::::::::::::::::::
214 GAGCGATTAAAGAACTACCCCAATTACACCCCGACATCATATTAA 263
69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  ::::::::::::::::::::
264 GGATGAGAGAAACACGGAGAGCAGCGGCTGATGACATGAGAGTCAAG 313
86 lueArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrrProGlyAl 102
  ::::::::::::::::::::
314 ACAAGTTAAATGCTTGGCCATCTCTGTGATGAACACAGTGGCTGAGTG 363
103 ArgLeuArgValThrGluGlyTrrAspGluAspGlyHisHisAlaGln 119
  ::::::::::::::::::::
364 AGCGTCGAGAGACGAGGCGTGGAGATGAGAGCCATCTTCAGAGGA 413
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTrsAspArg 136
  ::::::::::::::::::::
414 GTCTCTACATATGAGGTCGAGCAGTGCATCACACCGTCCGACCGG 463
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
  ::::::::::::::::::::
464 ACCGACCAAGTAGCAGTGTGCTGCTGCCCTGGCTGGAGGAGCGTTTC 513
153 AspTrrValTyrTrrGluSerArgAsnHisValHisValSerValLysAl 169
  ::::::::::::::::::::
514 GACGTGGGTCTCTATGATTCCAAGCTCACATCCATCTTCTGTGAAAGC 563
169 AspAsnSerLeuAlaValArgAlaGlyLy 179
  ::::::::::::::::::::
564 AGAGACTCGCTGGCGGCAAAATCCGCGCGC 594

```

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-08-460-900C-4

seq_documentation_block:

```

; Sequence 4, Application US/08460900C
; Patent No. 6155747
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; APPLICANT: Bumcrot, David A.
; APPLICANT: Marti-Gorostiza, Elisa
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,900C
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 4-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1313 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1311
; US-08-460-900C-4

```

```

alignment_scores:
  Quality: 740.00      Length: 177
  Ratio: 4.431
  Percent Similarity: 94.350      Percent Identity: 74.576

```

alignment_block:

US-08-900-220C-17_COPY_20_198 x US-08-460-900C-4 ..

Align seg 1/1 to: US-08-460-900C-4 from: 1 to: 1313

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaArgLy 19
  ::::::::::::::::::::
70 GCCGTGGGCCCCGAGGAGG...TTGGAAAGAGGCGGACCCCAAAA 116
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGln 36
  ::::::::::::::::::::
117 G...CTGACCCCTTACGCTACAGCAGTTATTTCCACAGTACCGAGA 163
36 rGTrLeuGlyAlaSerGlyProAlaGlyArgValAlaArgLySer 52
  ::::::::::::::::::::
164 AGACCCTAGGGGCGGCGGAGATGAGAGGAGATCAAGAAAGATCC 213
53 GluArgPheArgAspLeuValProAsnTyrTrsProAspIleIlePhe 69
  ::::::::::::::::::::
214 GAGCATTTAAGAACTACCCCAATTACAAACCCCGACATCATATTAA 263
69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  ::::::::::::::::::::
264 GGATGAGAGAAACACGGAGAGCAGCGGCTGATGATGAGAGTGCAGAG 313
86 lueArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrrProGlyAl 102
  ::::::::::::::::::::
314 ACAAGTTAAATGCTTGGCCATCTCTGTGATGAACACAGTGGCTGAGTG 363
103 ArgLeuArgValThrGluGlyTrrAspGluAspGlyHisHisAlaGln 119
  ::::::::::::::::::::
364 AGCGTCGAGAGACGAGGCGTGGAGATGAGAGCCATCTTCAGAGGA 413
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTrsAspArg 136
  ::::::::::::::::::::

```



```
APPLICATION NUMBER: US/08/460,900C
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 4-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1425
US-08-460-900C-6
```

```
alignment_scores:
Quality: 738.00      Length: 177
Ratio: 4.419         Gaps: 2
Percent Similarity: 94.350      Percent Identity: 74.011
```

```
alignment_block:
US-08-900-220c-17_copy_20_198 x US-08-460-900C-6 ..
```

```
Align seg 1/1 to: US-08-460-900C-6 from: 1 to: 1425
```

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgArgGlyTyrAlaArgly 19
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
67 GGGTGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 113
sglnleuValProleuLeuTyrGlnPheValProGlyValProGlyVal 36
| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
114 G...CTGACCCCTTTAGCCTACAGCAGTTTATCCCCAATGTGGCCGAGA 160
36 rghrleuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
161 AGACCTAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 210
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspPheLeuPhe 69
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
211 GAGCGATTTAAGGAACCTCACCCCAATTACACCCCGACATCATATTATTA 260
69 sspgluGlnAsnSerGlyAlaAspArgLeuMetPheGlnArgCysLysG 86
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
261 GGATGAGAAACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 310
86 luArgValAsnAlaLeuAlaIleAlaValMetAsnMetTyrProGlyVal 102
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
311 ACAAGTTGAACGCTTGGCATCTCGGTATGATACACAGTGGCGAGAGTGG 360
103 ArgLeuArgValThrGlnGlyTyrAspGlnAspGlyHisAlaGlnAs 119
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
361 AAACCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 410
119 pSerLeuHisTyrGlnGlyArgAlaLeuAspIleThrThrSerAspArg 136
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
411 GTCCTGCACTAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 460
```

```
136 spArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlyAlaGlyPhe 152
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
461 ACCGAGCAAGTACGCGCATGTGGCCCGCGGCGGCGGCGGCGGCGGCGGCTTC 510
153 AspTyrValTyrTyrGlnSerArgAsnHisValHisValSerValLysAl 169
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
511 GACTGGGTGTACTACGATCGCAAGCGCACATATTCACCTGCTCGGTGAAGC 560
169 aspAsnSerLeuAlaValArgAlaGlyLys 179
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
561 AGAAGAACTGCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 591
```

```
seq_name: /cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-10
```

```
seq_documentation_block:
```

```
: Sequence 10, Application US/08748591
```

```
: Patent No. 5759811
```

```
: GENERAL INFORMATION:
```

```
: APPLICANT: Epstein, Ervin
```

```
: APPLICANT: Hu, Zhilan
```

```
: APPLICANT: Bonifas, Jeanette
```

```
: TITLE OF INVENTION: Mutant Human Hedgehog Gene
```

```
: NUMBER OF SEQUENCES: 23
```

```
: CORRESPONDENCE ADDRESS:
```

```
: ADDRESS: Fish and Richardson
```

```
: STREET: 2200 Sand Hill Road
```

```
: CITY: Menlo Park
```

```
: STATE: CA
```

```
: COUNTRY: USA
```

```
: ZIP: 94025
```

```
: COMPUTER READABLE FORM:
```

```
: MEDIUM TYPE: Floppy disk
```

```
: COMPUTER: IBM PC compatible
```

```
: OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
: SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
: CURRENT APPLICATION DATA:
```

```
: APPLICATION NUMBER: US/08/748,591
```

```
: FILING DATE:
```

```
: CLASSIFICATION: 435
```

```
: ATTORNEY/AGENT INFORMATION:
```

```
: NAME: Sherwood, Pamela J
```

```
: REGISTRATION NUMBER: 36,677
```

```
: REFERENCE/DOCKET NUMBER: 06510/067001
```

```
: TELECOMMUNICATION INFORMATION:
```

```
: TELEPHONE: (415) 322-5070
```

```
: TELEFAX: (415) 854-0875
```

```
: INFORMATION FOR SEQ ID NO: 10:
```

```
: SEQUENCE CHARACTERISTICS:
```

```
: LENGTH: 1576 base pairs
```

```
: TYPE: nucleic acid
```

```
: STRANDEDNESS: single
```

```
: TOPOLOGY: linear
```

```
: MOLECULE TYPE: cDNA
```

```
: US-08-748-591-10
```

```
alignment_scores:
```

```
Quality: 734.00      Length: 177
```

```
Ratio: 4.395         Gaps: 2
```

```
Percent Similarity: 94.350      Percent Identity: 73.446
```

```
alignment_block:
```

```
US-08-900-220c-17_copy_20_198 x US-08-748-591-10 ..
```

```
Align seg 1/1 to: US-08-748-591-10 from: 1 to: 1576
```

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgArgGlyTyrAlaArgly 19
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
218 GCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 264
19 sglnleuValProleuLeuTyrGlnPheValProGlyValProGlyVal 36
| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
265 G...CTGACCCCTTTAGCCTACAGCAGTTTATCCCCAATGTGGCCGAGA 311
```

```

36  rgrhrleuglyalaserglyproalagluclyalargvalalargglyser 52
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 AGACCTGAGGCGGCGGAGGAGTATGAGAGGAGATCTCCAGAAACGCC 361
53  gluarprheargasplevalproasntryasnproaspillellepely 69
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 GAGCGATTTAAGAACTCACCCCAATTACACCCCGACATCATATTTAA 411
69  saspglucluasnseryalalaaspargleumethrqluarqcsylsg 86
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 GGATGAGAGAAACACGAGCGGAGCGCTGATGACTACAGAGGTGAAG 461
86  luargvalasnalaleualalalealavalmetasmettrproglyval 102
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 ACAAGTTGAAACGCTTTGGCCATCTCGGTGATAAACCAAGTGGCCGAGTG 511
103 Argleuargvalthrqluglytrpaspqluaspglyhishislaaglnas 119
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
512 AAATCGCGGTGACCGAGGCGTGGAGGAGAGATGCCACCTAGAGGA 561
119 pserleuhistyrglyalargalaleuaspilethrthserasparga 136
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
562 GTCTCTGACTACGAGGCGCGGAGTGCATCACACACGTCTGACCGCG 611
136 spargasnlystyrglyleuleualargleualavalglualaglyphe 152
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
612 ACCCCACCAAGTAGCGCATGCTGCCCGCCGCGGTGAGGCGCGCTTC 661
153 Asprtrvaltyrtyrgluserarfasnhsvalhishisvalservalysal 169
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
662 GACGGGTGACTACGAGTCCAGAGGACATATTCACATGCTCGTGAAAGC 711
169 aaspsnseryleualavalargalaglygly 179
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
712 AGAGAACTCGGTGGCGGCAAAATCGGAGGC 742

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-748-591-5
seq_documentation_block:
; Sequence 5, Application US/08748591
; Patent No. 5759811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Eryin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-748-591-5

alignment_scores:
  Quality: 732.00      Length: 177
  Ratio: 4.383        Gaps: 2
  Percent Similarity: 94.350  Percent Identity: 73.446

alignment_block:
US-08-900-220C-17_COPY_20_198 x US-08-748-591-5 ..
Align seq 1/1 to: US-08-748-591-5 from: 1 to: 1576

3  sercysgllyproglyalargglyprovalgllyarargargtrgryalargly 19
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
218 GCGTGGCGGACCGGCGGAGGGG...TTGCGGAGAGAGAGGACCCCAAAA 264
19  sglleuvalproleuethyrylsglnphevalproglyvalproglua 36
   | |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 G...CTACCCCTTACCTACAGCACTTATCTCCCAATGTGGCGGAGA 311
36  rgrhrleuglyalaserglyproalagluclyalargvalalargglyser 52
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 AGACCTGAGGCGGCGGAGGAGTATGAGAGGAGATCTCCAGAAACGCC 361
53  gluarprheargasplevalproasntryasnproaspillellepely 69
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 GAGCGATTTAAGAACTCACCCCAATTACACCCCGACATCATATTTAA 411
69  saspglucluasnseryalalaaspargleumethrqluarqcsylsg 86
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
412 GGATGAGAGAAACACGAGCGGAGCGCTGATGACTACAGAGGTGAAG 461
86  luargvalasnalaleualalalealavalmetasmettrproglyval 102
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
462 ACAAGTTGAAACGCTTTGGCCATCTCGGTGATAAACCAAGTGGCCGAGTG 511
103 Argleuargvalthrqluglytrpaspqluaspglyhishislaaglnas 119
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
512 AAATCGCGGTGACCGAGGCGTGGAGGAGAGATGCCACCTAGAGGA 561
119 pserleuhistyrglyalargalaleuaspilethrthserasparga 136
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
562 GTCTCTGACTACGAGGCGCGGAGTGCATCACACACGTCTGACCGCG 611
136 spargasnlystyrglyleuleualargleualavalglualaglyphe 152
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
612 ACCCGACCAAGTAGCGCATGCTGCCCGCCGCGGTGAGGCGCGCTTC 661
153 Asprtrvaltyrtyrgluserarfasnhsvalhishisvalservalysal 169
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
662 GACGGGTGACTACGAGTCCAGAGGACATATTCACATGCTCGTGAAAGC 711
712 AGAGAACTCGGTGGCGGCAAAATCGGAGGC 742

seq_name: /cgn2_6/ptodata/2/1na/6B_COMB.seq:US-08-460-900C-3
seq_documentation_block:
; Sequence 3, Application US/08460900C
; Patent No. 6165747
; GENERAL INFORMATION:
; APPLICANT: Ingam, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; APPLICANT: Bumcrot, David A.
; APPLICANT: Marti-Gorostiza, Elisa

```

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
 FILE OF INVENTION: Proteins and Uses Related Thereto
 NUMBER OF SEQUENCES: 62
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY, HOAG & ELIOT LLP
 STREET: One Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/460,900C
 FILING DATE: 5-JUNE-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/435,093
 FILING DATE: 4-MAY-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/356,060
 FILING DATE: 14-DEC-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/176,427
 FILING DATE: 30-DEC-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Vincent, Matthew P.
 REGISTRATION NUMBER: 36,709
 REFERENCE/DOCKET NUMBER: HMV-006.05
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 832-1000
 TELEFAX: (617) 832-7000
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1281 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1233
 OS-08-460-900C-3

alignment_scores:	
Quality:	729.50
Ratio:	4.421
Percent Similarity:	92.179
	Length: 179
	Gaps: 1
Percent Identity:	74.860

```
alignment_block:
US-08-900-220C-17_COPY_20_198 x US-08-460-900C-3  ..
```

Align seg 1/1 to: US-08-460-900C-3 from: 1 to: 1281

[illegible]

```

67 ephelysasgclugluasnserylalaaspargluethrpgluarG 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 CTTCAAGGAGGAGAGAACACGGGTGGCAGCCGCTCATGACCACGGCT 319
84 yslsgluarGvalAsnAlaleuAlaIleAlaValMetLasmethrpro 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 GCAGAGACCGCTCGAATCTCACTGGCGCATCTGTGATACACAGTGGCT 369
101 GlyValArgLeuArgValThrGluGlyTrpAspGluAspGluHisAla 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 GGTGTGAACCTGGCGGTGACCGAAGCCGGGATGAAGATGCCATCATCTC 419
117 aglnasrSerLeuHisThrGluGluGluAlaLeuAspIleThrTrpSerA 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 AGGAGGAGCTTTTCACATGAGGGCCGGCGGTGATATCACCACTCTAG 469
134 spArgAspArgAsnLysThrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCGTGACCCAAATTAATATGATGACTGCTGGCGCGCTTACCAATGCAGGCC 519
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVal 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GCGCTTGACACGGGATTTAGATGAGATCCAAAGCCCCACGTGATGCTCTGT 569
167 LysAlaAspAsnSerLeuAlaValArgAlaGlyTyr 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 CAAGTCTGACATTCGGCGCGCTGCCAACAACAGGTGGC 606

```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-176-427B-9

```
seq_documentation_block:
```

Patent No. 5789543

GENERAL INFORMATION:

APPLICANT: Ingham, Phillip W.

APPLICANT: McMahon, Andrew P.
TAB: Clifford J.

APPLICANT: Jadin, Clifford J.
TITLE OF INVENTION: Vertebral

TITLE OF INVENTION: Proteins and Uses Related Thereto

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & C

STREET: 60 State Street
CITY: Boston

STATE: MA
CITY: Boston

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy

COMPUTER: IBM PC comp

```

; OPERATING SYSTEM: PC
;
SOFTWARE: ASCII/text
;

```

```

; DO SWAP.  ASCII(LEAF)
; CURRENT APPLICATION DATA

```

APPLICATION NUMBER: 0

FILING DATE: 30-DEC-1

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION

```

; NAME: Vincent, Matthe
; REGISTRATION NUMBER:

```

REGISTRATION NUMBER.	REFERENCE/DOCKET NUMBER
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

TELECOMMUNICATION INFORMATION

TELEPHONE: (617) 227-

TELEFAX: (617) 227-59

; INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS
LENGTH: 1256 base pairs

LENGTH: 1250 base pairs
TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: CDNA

; FEATURE: NAME / REV. CDC

NAME/KEY:	CDS
LOCATION:	1 1257

LOCATION: T. 122 N.

US-08-176-427B-9

alignment_scores: Quality: 723.00 Length: 177
 Ratio: 4.409 Gaps: 2
Percent Similarity: 92.655 Percent Identity: 72.881

alignment_block:

US-08-900-220c-17_copy_20_198 x US-08-176-427B-9 ..

Align seg 1/1 to: US-08-176-427B-9 from: 1 to: 1256

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgGlyArgGly 19
  ::::::::::::::::::::::::::::|::::::::::::|
67 GCCGCGGCTCTGGCAGAGC...TACGCCAGAGAGACATCCGAAGAA 113
  |
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGln 36
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
114 G...CTGCACACTCTCGCCCTACAGCAGTTCATCTAATCTCCGCGAGA 160
  ::::::::::| ::::::::::| ::::::::::| ::::::::::|
36 rGThrLeuGlyAlaSerGlyProAlaGlnGlyArgValAlaArgGlySer 52
  ::::::::::| ::::::::::| ::::::::::| ::::::::::|
161 AGACCTTAGGGCCAGCGCGAGATACGAGCGCAAGATACGCCCATTCG 210
  |
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
211 GAGGATTTAAGAACTTACTCCAAATTACATCCGACATTAATCTTAA 260
  |
69 sAspGlnGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
261 GAGTGAAGAGAACCGGAGCGGAGCGAGCTCATGACACAGATGCAAG 310
  |
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  ::::::::::| ::::::::::| ::::::::::| ::::::::::|
311 ACAAGCTGAAGCTGGCCATCTCTAATGAACCACTGGCCAGGGGTT 360
  |
103 ArgLeuArgValThrGlnGlyTTPAspGluAspGlyHisHisAlaGlnAs 119
  ::::::::::| ::::::::::| ::::::::::| ::::::::::|
361 AAGCTGGCTGTGACAGAGAGGCTGGATGAGGACGCTCACCATTTTGA 410
  |
119 pSerLeuHisTyrGlnGlyArgAlaLeuAspIleThrThrSerAspArg 136
  ::::::::::| ::::::::::| ::::::::::| ::::::::::|
411 ATCAGCTCCACTACAGAGGAGAGAGCTGTTGATATTACCACTCTGACG 460
  |
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
461 ACAAGACCAATAATACGGGACACTGTCTGCTGAGCTGAGGATTT 510
  |
153 AspTrpValTyrTyrGlnSerArgAsnHisValHisValSerValLysAl 169
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
511 GACTGGGCTGTATTACGAGTCCAAAGCCACATTCATTCCTGTGCAAGC 560
  |
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
  | ::::::::::| ::::::::::| ::::::::::| ::::::::::|
561 AGAAATTTCGTTGCTGCAAAATCTGGGGC 591
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 11:01:20 ; Search time 57.94 Seconds

(without alignments)
362.102 Million cell updates/sec

Title: US-08-900-220c-17_COPY_20_198

Perfect score: 950
Sequence: 1 AOSCGPGGPGVGRRRYARKQ.....RNHHVSVKADNSLAVRAGC 179

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP_TREMBL_15:*
2: SP_Archaea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Mammal:*
8: SP_Misc:*
9: SP_Organelle:*
10: SP_Phage:*
11: SP_Plant:*
12: SP_Rodent:*
13: SP_Unclassified:*
14: SP_Vertebrate:*
15: SP_Virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	716	75.4	414	13	09W709 paralicthys
2	715	75.3	406	13	057567 notophthalm
3	712	74.9	434	13	057404 pleurodeles
4	664	69.9	129	11	09WMP6
5	640	67.4	415	5	017499
6	634	66.7	442	13	073803
7	618	65.1	139	6	09XSI6
8	600.5	63.2	471	5	09VQC4
9	599	63.1	138	13	09WGC1
10	563.5	59.3	410	5	061676
11	550	57.9	185	5	096699
12	525	55.3	150	13	09YGV7
13	523	55.1	161	11	09RI79
14	457	48.1	177	11	09WV29
15	375	39.5	88	13	09YCU3
16	367	38.6	87	5	09TX30
17	352	37.1	80	13	042441
18	340	35.8	227	5	09U526
19	331	34.8	119	13	042128

ALIGNMENTS

20	310	32.6	63	13	09IB14	09IB14 rana catesb
21	305	32.1	64	13	09PRF5	09PRF5 oryzias lat
22	228	24.0	49	5	09TX33	09TX33 hirudo medi
23	219	23.1	48	5	09TX31	09TX31 tritbolium c
24	171.5	18.1	48	5	09TX32	09TX32 strongyloce
25	90	9.5	127	2	09X758	09X758 streptomyc
26	89.5	9.4	490	8	P97956	P97956 unidentified
27	86	9.1	54	13	042233	042233 coturnix co
28	85.5	9.0	486	4	09UNF0	09UNF0 homo sapien
29	85	8.9	484	2	09LAV5	09LAV5 thermomonos
30	84.5	8.9	460	8	078521	078521 botrydium s
31	83.5	8.8	490	8	098947	098947 thalassiosl
32	81.5	8.6	342	14	058622	058622 royal farm
33	80.5	8.5	490	8	098945	098945 detonula co
34	79.5	8.4	471	5	023488	023488 caenorhabdi
35	78.5	8.3	476	8	046962	046962 rhizosoleni
36	78.5	8.3	680	2	052747	052747 ruminococcu
37	78.5	8.3	1193	4	092580	092580 homo sapien
38	78	8.2	675	10	09M0X5	09M0X5 arabidopsis
39	77.5	8.2	330	2	008347	008347 streptomyc
40	77.5	8.2	474	8	047285	047285 bumilleriop
41	77.5	8.2	488	8	09XPY4	09XPY4 lammarloco
42	77.5	8.2	523	14	09QBR1	09QBR1 canine dist
43	77.5	8.2	3670	2	09Z4X5	09Z4X5 streptomyc
44	77	8.1	1124	5	09U2C0	09U2C0 caenorhabdi
45	76.5	8.1	377	8	046979	046979 mallomonas

RESULT 1
ID 09W709 PRELIMINARY; PRT; 414 AA.
AC 09W709;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SONIC HEDGEHOG.
GN SHH.
OS Paralicthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Bothidae; Paralicthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99238226; PubMed=10223710;
RA Suzuki T., Ichiro O., Kurokawa T.;
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal
RT malformation in flounder (Paralicthys olivaceus) embryos.";
RL Dev. Growth Differ. 41:143-152(1999).
DR EMBL; AB029748; BAA82360.1; -.
DR HSSP; Q62226; 1VHH.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR INTERPRO: IPR001677; -.
DR PFAM; PF01079; Hint; 1.
DR PFAM; PF01085; HH_signal; 1.
DR PRINTS; PR00632; SONICHHOG.
DR SEQUENCE 414 AA; 45945 MM; 50607/BF3DB7C0DA3 CRC64;

Query Match 75.4%; Score 716; DB 13; Length 414;
Best Local Similarity 72.7%; Pred. No. 2.7e-58;
Matches 128; Conservative 28; Mismatches 18; Indels 2; Gaps 2;

Oy 4 CGPGPGVGRRRYARKQVPLLYKOFVGVGERTGASGPAEGRVRSERFDLYPVYN 63
||||| ||||| :| | | ||||| :| :| ||||| ||||| :| |||||
Db 24 CGPGPG-YGRRRPRK-LTPLAYKOFIPVNAEKTIAGSGRTGKITRNSRKELTPVYN 81

QY 64 PDIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 123
 DB 82 TDIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 141
 QY 124 EGRALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 179
 DB 142 EGRALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 197

RESULT 2
 ID 057567 PRELIMINARY: PRT: 406 AA.
 AC 057567

DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE HEDGEGHOG SEGMENT POLARITY HOMOLOG.
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
 OC Notophthalmus.
 NCBI_TaxID=8316;
 RX NCBI_TaxID=8316;
 RP SEQUENCE FROM N.A.
 RA Stark D.R., Gates P.B., Brookes J.P., Ferretti P.
 RL Dev. Dyn. 0:0-0(1998):1-1;
 DR EMBL: AF047466; AAC03108.1;
 DR HSSP: 062226; 1VNH.
 DR INTERPRO: IPR000320;
 DR INTERPRO: IPR001657;
 DR INTERPRO: IPR001767;
 DR INTERPRO: IPR002375;
 DR PFAM: PF01079; Hint: 1.
 DR PFAM: PF01085; HH_Signal: 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PROSITE: PS00103; PTR_PRR_PRR_TRANSFER: 1.
 DR PRODOM: PD003042; -; 1.
 SQ SEQUENCE 406 AA; 45072 MW; 5842CCAD5314DAD0 CRC64;

Query Match 75.3%; Score 715; DB 13; Length 406;
 Best Local Similarity 72.1%; Pred. No. 3.3e-58;
 Matches 129; Conservative 26; Mismatches 22; Indels 2; Gaps 1;
 QY 1 AOSCGRGPGVRRRARRKOLVPLLYKQFVGPVPTLGAAGPAGRGVARGSERFDLPVN 60
 DB 22 ALCCGGR--VIGRRRPPRLPLPSYKQFLPHVEKTLGASGVEGKIARNSRFEELTP 79
 QY 61 NYNDIIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 120
 DB 80 NYNDIIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 139
 QY 121 LAYEGALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 179
 DB 140 LAYEGALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 198

RESULT 3
 ID 057404 PRELIMINARY: PRT: 434 AA.
 AC 057404
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE SONIC HEDGEGHOG-RELATED PROTEIN.
 OS Pleurodeles waltlilii (Iberian ribbed newt).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
 OC Pleurodeles.
 NCBI_TaxID=8319;
 RX NCBI_TaxID=8319;
 RP SEQUENCE FROM N.A.

RA Caudit X., Nicolas S., Le Parco Y.
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF003532; AAB94412.1;
 DR HSSP: 062226; 1VNH.
 DR INTERPRO: IPR000320;
 DR INTERPRO: IPR001657;
 DR INTERPRO: IPR001767;
 DR PFAM: PF01079; Hint: 1.
 DR PFAM: PF01085; HH_Signal: 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; -; 1.
 SQ SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;

Query Match 74.9%; Score 712; DB 13; Length 434;
 Best Local Similarity 72.7%; Pred. No. 6.8e-58;
 Matches 128; Conservative 30; Mismatches 16; Indels 2; Gaps 2;
 QY 4 CGRGPGVRRRARRKOLVPLLYKQFVGPVPTLGAAGPAGRGVARGSERFDLPVN 63
 DB 27 CGRGPGVRRRARRKOLVPLLYKQFVGPVPTLGAAGPAGRGVARGSERFDLPVN 84
 QY 64 PDIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 123
 DB 85 PDIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 144
 QY 124 EGRALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 179
 DB 145 EGRALDITTSDBDRNRYGGLARLAVAGFDWYVESRNHVSVKADNSLAVRAG 200

RESULT 4
 ID 09WUP6 PRELIMINARY: PRT: 129 AA.
 AC 09WUP6
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE DESERT HEDGEGHOG PROTEIN (FRAGMENT).
 GN DHH.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RX NCBI_TaxID=10116;
 RP SEQUENCE FROM N.A.
 RA Garain P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;
 RT "Desert hedgehog in the rat."
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF148226; AAD31927.1;
 DR HSSP: 062226; 1VNH.
 DR INTERPRO: IPR000320;
 DR INTERPRO: IPR001657;
 DR PFAM: PF01085; HH_Signal: 1.
 DR PRINTS: PR00632; SONICHOG.
 FT NON_TER 1
 FT NON_TER 129
 SQ SEQUENCE 129 AA; 14578 MW; AB33509B3E751319 CRC64;

Query Match 69.9%; Score 664; DB 11; Length 129;
 Best Local Similarity 96.9%; Pred. No. 4.1e-54;
 Matches 123; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 2 OSGGPGRGVRRRARRKOLVPLLYKQFVGPVPTLGAAGPAGRGVARGSERFDLPVN 61
 DB 3 OSGGPGRGVRRRARRKOLVPLLYKQFVGPVPTLGAAGPAGRGVARGSERFDLPVN 62
 QY 62 YNDIIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 121
 DB 63 YNDIIFKEENSGADRLMTERCKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 122

QY 122 HYEGRAL 128
DB 123 HYEGRAL 129

RESULT 5

017499 PRELIMINARY; PRT; 415 AA.

AC 017499; 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
GN AMPHIH PROTEIN.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimeld S.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y13858; CA74169.1; -.
DR HSSP: Q62226; 1VHH.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR INTERPRO: IPR001767; -.
DR PFAM: PF01079; Hint; 1.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
DR PRODOM: PD003042; -; 1.
SQ SEQUENCE 415 AA; 46765 MW; 5E8D2BD2158D931 CRC64;

Query Match 67.4%; Score 640; DB 5; Length 415;
Best Local Similarity 66.5%; Pred. No. 2.9e-51;

Matches 119; Conservative 23; Mismatches 35; Indels 2; Gaps 2;

QY 1 AOSGCPGRCVGRRRYARKOLVPLLYKQFVPGVVERTLGASGPAEGRVARGSERFRDLPV 60
DB 25 SEACGPG-GAFGRRRHRPK-LTPFVYKQMPVASENFGASGLFNGRITRSEFRPHLKQ 82
QY 61 NYNDPIFKDEENSGADRLMTERCKERYNALAVNMMPGVRLVTEGDEDEGHHAODS 120
DB 83 NFNDIIFKDEEKTGADRFMTQRCCKLNLALISVMQWEGVKLRVTEGDEDEGFHTES 142
QY 121 LAYEGRALDITTSDRNRKYGILARLAVEAGFDWVYVESRHHVSVKADNSLAVRAGG 179
DB 143 LHTEGRAVDITTSDRNRKYGILARLAVEAGFDWVYVESKAIHCSVKAESDTTATOGG 201

RESULT 6
073803 PRELIMINARY; PRT; 442 AA.

AC 073803; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DE 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
GN FUGU HEDGEHOG.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Gellner K., Brenner S.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF056116; AAC34384.1; -.
DR HSSP: Q62226; 1VHH.
DR INTERPRO: IPR000320; -.

DR INTERPRO: IPR001767; -.
DR PFAM: PF01079; Hint; 1.
DR PFAM: PF01085; HH_signal; 1.
DR PRODOM: PD003042; -; 1.
SQ SEQUENCE 442 AA; 49286 MW; 1CB2B423B4B748E8 CRC64;

Query Match 66.7%; Score 634; DB 13; Length 442;
Best Local Similarity 67.4%; Pred. No. 1.1e-50;

Matches 120; Conservative 19; Mismatches 37; Indels 2; Gaps 2;

QY 2 QSCGPGPGVGRRRYARKOLVPLLYKQFVPGVVERTLGASGPAEGRVARGSERFRDLPV 61
DB 25 QSCGPGPG-YGIRTPRRK-LKAMYKQFPPLSENNGASGRAGKITRNSERFNELVCN 82
QY 62 NYNDPIFKDEENSGADRLMTERCKERYNALAVNMMPGVRLVTEGDEDEGHHAODS 121
DB 83 NYNDPIFKDEENSGADRLMTERCKERYNALAVNMMPGVRLVTEGDEDEGHHAODS 142
QY 122 HYEGRALDITTSDRNRKYGILARLAVEAGFDWVYVESRHHVSVKADNSLAVRAGG 179
DB 143 HYEGRALDITTSDRNRKYGILARLAVEAGFDWVYVESKAIHCSVKAESDTTATOGG 200

RESULT 7

09XS16 PRELIMINARY; PRT; 139 AA.

AC 09XS16; 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE SONIC HEDGEHOG (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TOOTH GERM;
RA Koyama E., Ivanoto M., Ohmori T., Kurisu K., Wu C., Ookura T.,
RA Bshiri M.M., Tucker T., Pacifici M.;
RT "Development of Stratum Intermedium and its Role as a Sonic Hedgehog-
Signaling Structure During Ootogenesis".
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF144100; AAD33926.1; -.
DR HSSP: Q62226; 1VHH.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHOG.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15961 MW; C83B0B96982046B1 CRC64;

Query Match 65.1%; Score 618; DB 6; Length 139;
Best Local Similarity 77.5%; Pred. No. 8.1e-50;

Matches 107; Conservative 22; Mismatches 9; Indels 0; Gaps 0;

QY 26 YKQFPGVPGVVERTLGASGPAEGRVARGSERFRDLPVNYNDPIFKDEENSGADRLMTERCK 85
DB 2 YKQFIPVNAEKTLLDASGRYEKITRNSERFELPNVNDPIFKDEENSGADRLMTORCK 61
QY 86 ERYNALAVNMMPGVRLVTEGDEDEGHHAODSLHYEGALDITTSDRNRKYGILAR 145
DB 62 DKLNALAVNMMPGVRLVTEGDEDEGHSESLHYEGRAVDITTSDRNRKYGILAR 121
QY 146 LAVEAGFDWVYVESRHHV 163
DB 122 LAVEAGFDWVYVESKAH 139

```
RESULT 8
Q9VCQ4 PRELIMINARY: PRT: 471 AA.
AC Q9VCQ4:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE HH GENE PRODUCT.
GN HH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE-20196006; PubMed-10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borlova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Dones M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li Z., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003742; AAF36102.1; -.
DR HSP: 002936; IAT0.
DR FLYBASE: FBgn0004644; hh.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR INTERPRO: IPR001767; -.
DR PFAM: PF01079; H1at; 1.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHHOG.
SQ SEQUENCE 471 AA; 52150 MW; 8ECD796A92FE7043 CRC64;
```

```
Query Match 63.2%; Score 600.5; DB 5; Length 471;
Best Local Similarity 62.6%; Pred. No. 1.5e-47;
Matches 112; Conservative 28; Mismatches 36; Indels 3; Gaps 2;
```

1 AASCGRGPGVGRRRARARQVPLLYKKQFVPGVPTLGAAGAEGRVARGSRFDLYP 60

```
Db 82 AASCGRGPG-IGRHR--ARNLYPLVLTQPTPNISEYTNASGPLEGVIARDSPKFKDLVP 138
QY 61 NYPNDIIFKDEENSGADRLMTTERCKEERVALAIAVNMMPGVLRTVEGMDEGHHAODS 120
Db 139 NNNRDLFLFDEBETGADRLMSRKCKLAVLAVSVNMNPGIRLVLTVESMDEYHHGQDS 198
QY 121 LHYEGALDITTSDDRNRKYGGLARLAVAGFDWYVESRNMHVSVKADNSLAVRAG 179
Db 199 LHYEGRAVLIATSDRQSGKXGMLARLAVAGFDWVSYSRRHLYCSVKSDSSISHVHG 257

RESULT 9
Q9W6C1 PRELIMINARY: PRT: 138 AA.
AC Q9W6C1:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE SONIC HEDGEHOG PROTEIN (FRAGMENT).
GN SHH.
OS Eleutherodactylus coqui.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Eleutherodactylus.
OC NCBI_TaxId=57060;
RN [1]
RP SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui.";
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF113403; AAD23436.1; -.
DR HSP: 062226; 1VNH.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHHOG.
FT NON_TER 1
FT NON_TER 138
FT TER 138
SQ SEQUENCE 138 AA; 15751 MW; FFA156A17F4681F0 CRC64;
```

```
Query Match 63.1%; Score 599; DB 13; Length 138;
Best Local Similarity 75.4%; Pred. No. 4.6e-48;
Matches 104; Conservative 23; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 30 VPGVPERTLGASGPARGARQSERPRDLVPYNNPDITFKDENSGADRLMTTERCKERYN 89
Db 1 IPVVAEKTYLGASGRGEGKITRNSERKELTPVNSDIIFKDEENTVAADRLMTQORCKDKLN 60
QY 90 ALAIAVNMNMPGVRLLRTVEGMDEGHHAODSLHYEGRALDITTSDDRNRKYGGLARLAVE 149
Db 61 ALAIAVNMNMPGVRLLRTVEGMDEGHHAODSLHYEGRAVDITTSDDRNRKYGGLARLAVE 120
```

```
QY 150 AGFDWYVESRNMHVSV 167
Db 121 AGFDWYVESKMHICSV 138
```

```
RESULT 10
061676 PRELIMINARY: PRT: 410 AA.
AC 061676:
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE HEDEHOG.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
```


OX NCBI_TaxID=7654;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hertzel P.L., McClay D.R.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF059606; AAC15065.1; -
 DR HSSP: Q62226; 1VHH.
 DR INTERPRO: IPR000320; -
 DR INTERPRO: IPR001657; -
 DR INTERPRO: IPR001767; -
 DR PFAM: PF01079; Hntf; 1.
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR PRODOM: PD003042; -; 1.
 SO SEQUENCE 410 AA; 46071 MW; 875CF8989DA6FDBED CRC64;

 Query Match 59.3%; Score 563.5; DB 5; Length 410;
 Best Local Similarity 60.7%; Pred. No. 3.3e-44;
 Matches 108; Conservative 25; Mismatches 42; Indels 3; Gaps 2;

 OY 2 QSCGPGRRYARRKQVLPVLYKQFVPGVPRTLGASGAPGRRVARGSEFRDLYPN 61
 DB 25 QACHPGRS--GKTSRRRNRPTLOYKQVRISSDTFGASGPPGRIDRDDEFSKLSPN 82
 OY 62 YNPDIIFKDEENSGADRLMTERRCKERYNALAIAYNMMPGVRLEVTGWDGHHADSL 121
 DB 83 NNDIVFKDEEGCAGDRLMTQRCCKDLNTLAISYNNMPGKIKLAVFVAMWDEQPNV-EPL 141
 OY 122 HYERADITTSDDRRKYGGLARLAVAGFDWYVESRNIVHVSVAADNSLAVRAG 179
 DB 142 HAEGRADITTSDDRRKYGGLARLAVAGFDWYVESKAWHCSVSESAANSG 199

 RESULT 11
 096699 PRELIMINARY; PRT; 185 AA.
 AC 096699;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE HEDGEHOG PROTEIN (FRAGMENT).
 GN HH.
 OS Junonia coenia (Peacock butterfly) (Precis coenia).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
 OX NCBI_TaxID=39708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
 RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
 RT "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
 evolution.";
 RL Science 0:0-0(1999).
 DR EMBL: AF117742; AAD08931.1; -
 DR HSSP: Q62226; 1VHH.
 DR INTERPRO: IPR00320; -
 DR INTERPRO: IPR001657; -
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR NON_TER 1
 FT NON_TER 185
 SO SEQUENCE 185 AA; 20745 MW; E9B9F4C2FAA662 CRC64;

 Query Match 57.9%; Score 550; DB 5; Length 185;
 Best Local Similarity 69.8%; Pred. No. 2.2e-43;
 Matches 97; Conservative 22; Mismatches 20; Indels 0; Gaps 0;

 OY 40 ASGAEGRVARGSRPDLVYNNPDIFKDEENSGADRLMTERRCKERYNALAIAYNM 99
 DB 3 ASGPEGRITVDDKFRDLVYNNPDIDFKDEEGADRLMTQRCCKEKLNTLAISVNM 62

OY 100 PGVRLRVTEGDEGHHADSLHGEALDITTSDDRRNKYGGLARLAVAGFDWYVES 159
 DB 63 PGVRLRVTEGDEENSHLNSLHGEAGVADLTTSDDRRSKNGMLARLAVAGFDWYVEEN 122

 OY 160 RNHVHVSFKADNSLAVRAG 178
 DB 123 RSYIHCYKTESVGTGAG 141

 RESULT 12
 096699 PRELIMINARY; PRT; 150 AA.
 AC 096699;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE SONIC HEDGEHOG (FRAGMENT).
 OS Ambystoma mexicanum (Axolotl).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomidae;
 OX NCBI_TaxID=8296;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Torok M.A., Izpizua-Belmonte J.C., Gardiner D.M., Bryant S.V.;
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF031480; AAD18128.1; -
 DR HSSP: Q62226; 1VHH.
 DR INTERPRO: IPR000320; -
 DR INTERPRO: IPR001657; -
 DR PFAM: PF01085; HH_signal; 1.
 DR PRINTS: PR00632; SONICHOG.
 DR NON_TER 1
 FT NON_TER 150
 SO SEQUENCE 150 AA; 16599 MW; 986F5037A69A8AC CRC64;

 Query Match 55.3%; Score 525; DB 13; Length 150;
 Best Local Similarity 79.1%; Pred. No. 3.5e-41;
 Matches 91; Conservative 20; Mismatches 4; Indels 0; Gaps 0;

 OY 65 DIFKDEENSGADRLMTERRCKERYNALAIAYNMMPGVRLEVTGWDGHHADSLH 124
 DB 1 DIFKDEENSGADRLMTQRCCKDLNTLAISYNNMPGKIKLAVFVAMWDEQPNV-EPL 60

 OY 125 GRADITTSDDRRKYGGLARLAVAGFDWYVESRNIVHVSVAADNSLAVRAG 179
 DB 61 GRAVDITTSDDRRSKYGGLARLAVAGFDWYVESKAHHCVAENSAKSG 115

 RESULT 13
 096699 PRELIMINARY; PRT; 161 AA.
 AC 096699;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 15, Last annotation update)
 DE SONIC HEDGEHOG PROTEIN (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-SPRAGUE-DAWLEY; TISSUE-KIDNEY;
 RA Garces P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;
 RT "Sonic hedgehog in the rat.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF162915; AAD45373.1; -
 DR HSSP: Q62226; 1VHH.
 DR INTERPRO: IPR000320; -
 DR INTERPRO: IPR001657; -

THIS PAGE BLANK (USPTO)

OM of: US-08-900-220c-17_COPY_20_198 to: EST:* out_format : pfs
Date: Apr 22, 2001 11:43 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+g2n.model -DEV=xlp
-O/cgn2_1/USP001/US08900220/runat_19042001_104047_21657/app-query.fasta.1.253
-B-EST -QFMT=fastap -SUFRT=1st -GAPOP=12.000 -GAPEXT=4.500
-MINMATCH=0.100 -LOOPT=0.000 -LOOPTXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-START=1 -MATRIX=blissum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=pcy -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -DOCCALIGN=200 -OUTFMT=pfcs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US08900220 @cgn1_1_2790 -NCPU=6
-ICPU=3 -LONGLOC -NO_XLPRX -WAIT -THREADS=1

Search information block:
Query: US-08-900-220c-17_COPY_20_198
Query length: 179
Database: EST:*
Database sequences: 9623517
Database length: 73081774
Search time (sec): 1170.070000

Score_list:	Strd Orig	ZScore	EScore	Len	Documentation
Sequence					
gb_est14:AF281007	568.00	136.73	3.2e-54	519	AF281007 Cyprinus carpio
gb_est14:AF281008	495.50	990.06	4.7e-46	514	AF281008 Cyprinus carpio
gb_est14:AF281009	437.00	877.18	9.1e-40	299	AF281009 Cyprinus carpio
gb_est18:BF601365	415.00	828.71	4.6e-37	441	BF601365 Cyprinus carpio
gb_est13:AF1645932	406.00	808.49	6.1e-36	537	AF1645932 Cyprinus carpio
gb_est13:AF166359	391.00	779.63	2.5e-34	463	AF166359 Cyprinus carpio
gb_est13:AF166359	352.00	702.03	5.2e-30	406	AF166359 Cyprinus carpio
gb_est14:AF201281	318.50	628.12	6.9e-26	748	AF201281 Cyprinus carpio
gb_est14:AF201255	313.50	619.73	2.0e-25	625	AF201255 Cyprinus carpio
gb_gss31:CN504F35	308.00	604.60	1.4e-24	926	AF281007 Cyprinus carpio
gb_gss32:CN505M56	222.50	432.80	5.4e-12	819	AF281007 Cyprinus carpio
gb_est18:AF603741	195.00	387.65	1.7e-12	291	AF603741 Cyprinus carpio
gb_est15:AF1015155	180.00	352.66	1.5e-10	459	AF1015155 Cyprinus carpio
gb_est13:AF604922	165.00	320.37	9.4e-09	555	AF604922 Cyprinus carpio
gb_est14:BF030536	160.00	318.80	1.2e-08	239	BF030536 Cyprinus carpio
gb_est14:AF076750	139.00	256.19	3.5e-05	288	AF076750 Cyprinus carpio
gb_est16:AF1124084	113.50	218.33	0.0045	447	AF1124084 Cyprinus carpio
gb_gss29:CN501SEN	94.00	173.03	1.51	793	AF1124084 Cyprinus carpio
gb_gss29:AF0889184	91.50	164.95	4.28	1069	AF0889184 Cyprinus carpio
gb_est16:BE615322	90.50	164.12	4.75	950	BE615322 Cyprinus carpio
gb_est17:AA430730	89.50	170.17	2.19	429	AA430730 Cyprinus carpio
gb_est18:BE962904	88.00	161.90	6.32	718	BE962904 Cyprinus carpio
gb_est13:AF136206	87.00	163.78	4.97	489	AF136206 Cyprinus carpio
gb_est13:BE410712	87.00	158.91	9.27	790	BE410712 Cyprinus carpio
gb_est13:AF1521451	84.50	153.31	19.01	833	AF1521451 Cyprinus carpio
gb_est13:BF1138455	84.00	151.36	24.43	914	BF1138455 Cyprinus carpio
gb_est15:BF303653	84.00	145.08	11.12	1696	BF303653 Cyprinus carpio
gb_est12:BF805523	83.50	157.50	5.45	452	BF805523 Cyprinus carpio
gb_est17:BG170046	83.50	150.03	28.97	943	BG170046 Cyprinus carpio
gb_est18:BF533134	82.50	148.39	16.03	908	BF533134 Cyprinus carpio
gb_est10:AF0162172	82.00	154.64	3.75	444	AF0162172 Cyprinus carpio
gb_est19:BF973856	82.00	148.65	34.58	801	BF973856 Cyprinus carpio
gb_est10:BE905658	82.00	144.95	55.57	1153	BE905658 Cyprinus carpio
gb_est13:AA884048	81.50	155.43	14.50	372	AA884048 Cyprinus carpio
gb_est13:BF106035	81.50	148.79	33.95	715	BF106035 Cyprinus carpio
gb_est19:AF035880	81.50	148.03	37.46	771	AF035880 Cyprinus carpio
gb_est11:BE965326	81.50	139.55	111.06	1776	BE965326 Cyprinus carpio
gb_est18:C98572	81.00	152.59	20.85	445	C98572 Cyprinus carpio
gb_est13:BF146039	81.00	149.89	29.50	581	BF146039 Cyprinus carpio
gb_est17:BE214706	81.00	149.03	32.92	632	BE214706 Cyprinus carpio
gb_est13:BE414016	81.00	148.08	37.19	694	BE414016 Cyprinus carpio

gb_est14:BF236001	+	81.00	147.53	39.93	733	BF236001 602025687F1 NCI_CGA
gb_est15:BF317393	+	81.00	146.96	42.94	775	BF317393 601903730F1 NIH_MGC
gb_est18:BF536960	+	81.00	146.41	46.07	818	BF536960 602048859F1 NCI_CGA
gb_est10:W97573	+	80.50	151.74	23.25	438	W97573 m97570.r1 Soares mou
seq_name: gb_est14:AF281007						
seq_documentation_block:						
LOCUS	AF281007	519 bp	mRNA	EST	13-NOV-2000	
DEFINITION	AF281007 Cyprinus carpio library (Hall T)					Cyprinus carpio cDNA
ACCESSION	AF281007					
VERSION	AF281007.1	GI:11141065				
KEYWORDS	EST.					
SOURCE	common carp.					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.					
REFERENCE	1 (bases 1 to 519)					
AUTHORS	Cole, N.J., Hall, T.E. and Johnston, I.A.					
TITLE	mRNA sequence					
JOURNAL	Unpublished (2000)					
COMMENT	Contact: Hall T Gatty Marine Laboratory University of St. Andrews East Sands, St. Andrews, Fife KY16 8LB, UK Email: thomasehall@hotmail.com.					
FEATURES						
source	1..519					
	/organism="Cyprinus carpio"					
	/db_xref="taxon:7962"					
	/clone_id="Cyprinus carpio library (Hall T)"					
BASE COUNT	154 a 114 c 138 g 106 t					7 others
ORIGIN						
alignment_scores:						
Quality:	568.00					Length: 134
Ratio:	4.403					Gaps: 0
Percent Similarity:	96.269					Percent Identity: 71.642
alignment_block:						
US-08-900-220c-17_COPY_20_198 x AF281007						
Align seg 1/1 to: AF281007 from: 1 to: 519						
45 GI:11141065	AlaArgGlySerGluArgPheArgPleuValProAs	61				
8 GAGGGCAAGATCCGCGCATTCGAGAGATTAAAGAACTTACTCCAAA		57				
61 nTyrAsnProAspLeuIlePheLysAspGluGluAsnSerGlyAlaAspA		78				
58 TTACATATCCGATATATTCCTTAAGATGAGGAGAACTGGCGGGGACA		107				
78 rGLeuMetThrGluArgCysLysGluArgValAsnAlaLeuAlaIleAla		94				
108 GACATCATGACACAGATGCAAGAAAGCAAGCTGACACTGCGCATCTCT		157				
95 ValMetAsnMetTrpProGlyValArgLeuAlaValThrGluGlyTrpAs		111				
158 GTAAAGAACCACTGGCGAGGGGTAAAGCTGTGACAGAGGCTGGGA		207				
111 pGluAspGlyHisHisAlaGlnAspSerLeuHisIleGlyGluGlyAlaAla		128				
208 TGAGAGACGCTACCACTTTTGAAGATCTACTCCATCAAGAGGAGGCTG		257				
128 euAspIleThrThrSerAspArgAsnLysTyrGlyLeuLeuAla		144				
258 TTGATATATCAACCACTGATCAAGAGAGCAATATGAGCACTCTCT		307				
145 ArgLeuAlaValAlaGluAlaGlyPheAspTrpValTyrTyrGluSerArgAs		161				

```

308 CGGTACCTGCGAGCTGATTTGACTGGCTCTATTACAGATNCAAGC 357
161 nhlsvalhisvalserVallysalaspasnsSerleualavalargalag 178
358 CCACATTCATGCTGCTGTCAAGCAGAAATTCGGTTGCGGCAATCTG 407
178 ly 178
408 GA 409
seq_name: gb_est14:AF281008

seq_documentation block:
LOCUS AF281008 514 bp mRNA EST 13-NOV-2000
DEFINITION AF281008 Clupea harengus library (Hall T) Clupea harengus cDNA
similar to sonic hedgehog, mRNA sequence.
ACCESSION AF281008
VERSION AF281008.1 GI:11141066
KEYWORDS EST
SOURCE Atlantic herring.
ORGANISM Clupea harengus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
Clupeinae; Clupea.
1 (bases 1 to 514)
Cole,N.J., Hall,T.E. and Johnston,I.A.
AUTHORS mRNA sequence
JOURNAL Unpublished (2000)
COMMENT Contact: Hall T
Gatty Marine Laboratory
University of St. Andrews
East Sands, St. Andrews, Fife KY16 8LB, UK
Email: thomasehall@hotmail.com.
FEATURES
source 1..514
/organism="Clupea harengus"
/db_xref="taxon:7950"
/clone_lib="Clupea harengus library (Hall T)"
BASE COUNT 150 a 116 c 144 g 99 t 5 others
ORIGIN

alignment_scores:
Quality: 495.50 Length: 136
Ratio: 4.061 Gaps: 2
Percent Similarity: 89.706 Percent Identity: 67.647

alignment_block:
US-08-900-220c-17_COPY_20_198 x AF281008 ..
Align seg 1/1 to: AF281008 from: 1 to: 514
45 GUGUAGYARValAlaArglySerGluArg.PheArgAspLeuValProA 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
8 GAGGCGAAGATCAACGGCGCATTTNGCGAGAGTTTAAAAAACCTNCTC...C 54
61 snTyRasnProAspIleIlePheIlyAspGluGluAsnSerGlyAlaAsp 77
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
55 AAATNCAATCCGACATCATCTTTAAGGATGAGGAGACAGCGGCGCGAC 104
78 ARIleuMetHrthGluArgCysLySGluArgValAsnAlaLeuAlaIleA 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
105 AGACTCATGACACAGATGACAAAGACAGCTGAACCTGACCTGCATCTC 154
94 aValMetAsnMetTrpProGlyValArgLeuArgValThGluGlyTrpA 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
155 TGTATATGAACACAGCGCGGCGGTCACACTGCGTGACACAGAGACTGCG 204
111 spGluAspGlyHisHisAlaGlnAspSerLeuHisTyrgluGlyArgAla 127
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
205 GTGAGGAGCGGTACACATTTTGAAGAAATCACTCCACTGAGGAGGAGGCT 254
128 LeuAspIleHrthSerAspArgAspArgAsnLysTyrglyLeuLeuAl 144

```

```

255 GTTGATATATCACACACCTCTGATCGACAGACAGAGATATGTGACACTCTC 304
144 argLeuAlaValGluAlaGlyPheAspTrpValIlyTyrglySerArgA 161
305 TCGGTTAGCTGTGGAAGCTGGATTTCGCTGCTATATACAGATCCAAAG 354
161 snHlsValHisValSerVallysalaspasnsSerleualavalargal 177
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
355 CCCAACACCATTCCTCTGTCAAGCAGAAATTCGGTTGCGGCAATCT 404
178 GlyGly 179
405 GGAAGC 410
seq_name: gb_est42:AW144802

seq_documentation block:
LOCUS AW144802 299 bp mRNA EST 30-OCT-1999
DEFINITION EST291875 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone
RG10G52.5' end similar to Indian hedgehog protein, mRNA sequence.
ACCESSION AW144802
VERSION AW144802.1 GI:6161619
KEYWORDS EST
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 299)
Lee,N.H., Glodek,A., Chandra,I., Mason,T.M., Quackenbush,J.,
Kerlavage,A.R. and Adams,M.D.
AUTHORS Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT Contract: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@ligr.org
For clone availability, additional sequence and expression
information related to this EST please check the TIGR Rat Gene
Index (http://www.tigr.org/tdb/rgi/rgi.html). To order a clone
contact the ATCC (http://www.atcc.org/atcc.html).
Seq primer: M13 Reverse.
FEATURES
source 1..299
/organism="Rattus sp."
/db_xref="taxon:10118"
/clone="RG10G52"
/clone_lib="Normalized rat embryo, Bento Soares"
/dev_stage="embryo 8, 12, 18 dpc"
/note="Vector: pT73Pac; Site_1: EORI; Site_2: NotI"
BASE COUNT 78 a 92 c 80 g 49 t
ORIGIN

alignment_scores:
Quality: 437.00 Length: 99
Ratio: 4.600 Gaps: 0
Percent Similarity: 95.960 Percent Identity: 77.778

alignment_block:
US-08-900-220c-17_COPY_20_198 x AW144802 ..
Align seg 1/1 to: AW144802 from: 1 to: 299
41 SerGlyProAlaGluGlyArgValAlaArgGlySerGluArgPheArgAs 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 AGCGGGCGGTACTAAGCAAAATTCGCGGCGAGCTGTGAGGCTTCAAGGA 51
57 pleuValProAsnTyRasnProAspIleIlePheIlyAspGluGluAsn 74

```


TGTACCAATCTGAAGTGGAGCCGCCGTTTTTTTTTTTTTTTTTTTTT
 31; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M. Fatima Bonaldo.

alignment_scores: Quality: 406.00 Length: 82
 Ratio: 5.012 Gaps: 0
 Percent Similarity: 98.780 Percent Identity: 93.902

alignment_block:
 US-08-900-220C-17_COPY_20_198 x A1645932 ..
 Align seg 1/1 to: A1645932 from: 1 to: 537

1 AlagInSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
 |||||||
 292 GCCAGAGAGCTGGGGCGGGCGGAGACCGGTTGCCGCGCGGTATGT 341
 17 aArgGlyGlnLeuValProLeuLeuTyrGlnPheValProGlyValP 34
 :|||||
 342 GCGCAACCACTTGCTGCTGTATCAAGCACTTGTCCCGCATATGC 391
 |||||||
 34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
 |||||||
 392 CCGAGCGGACCTTGTCGCGAGTGCGCCAGCGGAGGAGGATACAGG 441
 |||||||
 51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle 67
 |||||||
 442 GGGTCGAGGAGCTTCCGGACCTGCTACCACTACAACCCCGCATAT 491
 |||||||
 67 ephelyAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGlu 82
 |||||||
 492 CTTCACGATGAGAGAACAGCGCGCAGACCGCTGATGACAGAG 537
 seq_name: gb_est23:A1666359

seq_documentation_block:
 LOCUS A1666359 463 bp mRNA EST 12-MAY-1999
 DEFINITION mul1906.x1 Soares.thymus_2NDMT Mus musculus cDNA clone IMAGE:639130
 3' similar to YR:061488 Q61488 DESERT HEDGHEOG HOMOLOG PRECURSOR ;,
 mRNA sequence.

ACCESSION A1666359
 VERSION A1666359.1 GI:4804713
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 463)
 NCI-GAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: mul1906.y1
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Email: Robert.Strausberg@nih.gov
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 Possible reversed clone: similarity on wrong strand
 High quality sequence stop: 261.
 Location/Qualifiers
 1..463
 /organism="Mus musculus"

/strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:639130"
 /clone_lib="Soares_thymus_2NDMT"
 /sex="male"
 /tissue_type="thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer 15'
 TGTACCAATCTGAAGTGGAGCGCGCGGTTTTTTTTTTTTTTTTTT
 31; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M. Fatima Bonaldo."

alignment_scores: Quality: 391.00 Length: 82
 Ratio: 4.949 Gaps: 0
 Percent Similarity: 96.341 Percent Identity: 91.463

alignment_block:
 US-08-900-220C-17_COPY_20_198 x A1666359/rev ..
 Align seg 1/1 to reverse of: A1666359 from: 1 to: 463

1 AlagInSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
 |||||||
 406 GCCAGAGAGCTGGGGCGGGCGGAGACCGGTTGCCGCGCGGTATGT 357
 17 aArgGlyGlnLeuValProLeuLeuTyrGlnPheValProGlyValP 34
 :|||||
 356 GCGCAACCACTTGCTGCTGTATCAAGCACTTGTCCCGCATATGC 307
 |||||||
 34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
 |||||||
 306 CCGAGCGGACCTTGTCGCGAGTGCGCCAGCGGAGGAGGATACAGG 257
 |||||||
 51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle 67
 |||||||
 256 GGGTCGAGGAGCTTCCGGACCTGCTACCACTACAACCCCGCATAT 207
 |||||||
 67 ephelyAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGlu 82
 |||||||
 206 CTTCACGATGAGAGAACAGCGCGCAGACCGCTGATGACAGAG 161
 seq_name: gb_est27:A1958076

seq_documentation_block:
 LOCUS A1958076 406 bp mRNA EST 20-AUG-1999
 DEFINITION fc90e01.y1 zebrafish washu MPING EST Danio rerio cDNA 5' similar to
 TR:092008 Q92008 VHH-1 PRECURSOR ;, mRNA sequence.
 ACCESSION A1958076
 VERSION A1958076.1 GI:5750785
 KEYWORDS EST.
 SOURCE zebrafish.
 ORGANISM Danio rerio

REFERENCE
 AUTHORS
 Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,
 S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,
 K., Stepien,M., Theisling,B., Allen,M., Bowers,Y., Person,B.,
 Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
 Kohn,S., Shiu,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
 and Wilson,R.

TITLE Washu Zebrafish EST Project 1998
JOURNAL
COMMENT

Unpublished: (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratfish@wustl.edu
CDNA Library Preparation: Matthew Clark. CDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
Resource Center (umprimardatenbank, Berlin, Germany (web address:
www.rzpd.de)

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: T3 ET from AmerSham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

SOURCE

1. 406
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="Zebrafish Washu MPIMC EST"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield
stage embryos"
/lab_host="XLI-blue MRP"
/note="Vector: pSPORT1. Site_1: NotI. Site_2: SalI. 1st
strand cDNA was primed with a Not I - oligo(dT)15 primer
[5'-pGACTGATCTCTACATCGGAGCGCCGCCCTTTTCTTTT3'];
double-stranded cDNA was ligated to Sal I adaptors (BRL),
digested with Not I and cloned into the Not I and Sal I
sites of the pSPORT1 vector (BRL). Library was constructed
by Matthew Clark (Lehrbach lab, ICRF, London and Max Planck
Institut fuer Molekulare Genetik, Berlin). cDNAs for EST
analysis were selected following oligonucleotide
hybridization fingerprinting of arrayed clones from
zebrafish late somitogenesis (26 ss), adult liver or
embryonic shield stage (5-6 h) libraries. Fingerprint
data were used to computationally cluster cDNAs, and a
single cDNA from each cluster was chosen for sequencing.
In some cases multiple members of the same cluster were
sequenced to assess clustering parameters or single clones
were sequenced additional times to assess quality
control."

BASE COUNT 106 a 95 c 113 g 92 t
ORIGIN

alignment_scores:
Quality: 352.00 Length: 116
Ratio: 3.705 Gaps: 2
Percent Similarity: 81.897 Percent Identity: 59.483

alignment_block:

US-08-900-220c-17_COPY_20_198 x A1958076 ..

Align seg 1/1 to: A1958076 from: 1 to: 406

```

3  SerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaArgly 19
   :::::::::::::::::::: 111 11111111111111111111
65 GCCTCGCGGCTGCTTAGGC...TACGGCAGACTAAGACATCG...AA 108
   :::::::::::::::::::: 111 11111111111111111111
19 scInleuValProleuLeuTyrlysgInpheValProGlyValProGluA 36
   |:::::::::::::::::: 11111111111111111111111111
109 GATGCGACACTCTCTACAGGGTTCATACCTAATGTGAGAGCA 158
   :::::::::::::::::::: 111 11111111111111111111
36 rgtThrleuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
   :::::::::::::::::::: 111 11111111111111111111
159 AACACCTATAGGCGCAGCGCAGATACAGGCATGATACGCCCATTTAG 208

```

```

53 GluArgpheArgAspLeuValProAsnTyrAsnProAspIleIlephely 69
   |:::::::::::::::::: 111 11111111111111111111
209 GAGAGATTGAAGAACTTACTCCATATTACATGCGCAGCATATCTTTA 258
   :::::::::::::::::::: 111 11111111111111111111
69 sAspGluGuaSerserGlyAlaAspArgleuMetThrGluArgCyslysg 86
   |:::::::::::::::::: 11111111111111111111111111
259 GGATAGAGCAACACGCTAGCGAGACGCTCATGACACGAGATCCAAAG 308
   :::::::::::::::::::: 111 11111111111111111111
86 luArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
   :::::::::::::::::::: 111 11111111111111111111
309 ACAAGCTGATCTCGCTGCGCATCTATGTATGACACACGCGCAGTCGTT 358
   :::::::::::::::::::: 111 11111111111111111111
103 ArgLeuArgValThrGluGlyTrrpAspGluAspGlyHisAlaGln 118
   :::::::::::::::::::: 111 11111111111111111111
359 AAGCTGCGGTGACAGACGCTGAGATGAGAGCGCGCACCATCTTGAA 406
   :::::::::::::::::::: 111 11111111111111111111

```

seq_name: qb_est14:AF201281

seq_documentation_block:

LOCUS AF201281 748 bp mRNA EST 13-NOV-2000
DEFINITION AF201281 Botryllus schlosseri EST Botryllus schlosseri cDNA, mRNA
sequence.

ACCESSION AF201281 GI:11141061

VERSION AF201281

KEYWORDS EST.

SOURCE Botryllus schlosseri.

ORGANISM Botryllus schlosseri.

REFERENCE 1 (bases 1 to 748)

AUTHORS Meyer, K.A., Leonard, P.M., Rosenberg, G.H., Miller, R.D. and Loker, E.S.

TITLE Characterization of ESTs from the Colonial Tunicate, Botryllus schlosseri

JOURNAL Unpublished (2000)

COMMENT Contact: Loker ES

Parasitology

Biology, University of New Mexico

Castetter Hall, Albuquerque, NM 87131, USA

Tel: 505 277 5508

Fax: 505 277 0304

Email: esloker@unm.edu.

FEATURES

source 1. 748

Location/Qualifiers

/organism="Botryllus schlosseri"

/db_xref="taxon:30301"

/clone_lib="Botryllus schlosseri EST"

BASE COUNT 205 a 200 c 184 g 159 t

ORIGIN

alignment_scores:
Quality: 318.50 Length: 162
Ratio: 2.611 Gaps: 5
Percent Similarity: 75.309 Percent Identity: 41.975

alignment_block:

US-08-900-220c-17_COPY_20_198 x AF201281 ..

Align seg 1/1 to: AF201281 from: 1 to: 748

```

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgTyrAla 17
   |:::::::::::::::::: 111 11111111111111111111
148 GCGGAAGCTGC...CGGGCCCGCGGCGCGGCTTTTCGCACGACACCC 194
   :::::::::::::::::::: 111 11111111111111111111
17 bArglysgInleuValProleuLeuTyrlysgInpheValProGlyValP 34
   :::::::::::::::::::: 111 11111111111111111111
195 TCACAAGACGCAAGTGGCGCTGATATCCGCGCAGCATATTCACAACTGT 244
   :::::::::::::::::::: 111 11111111111111111111
34 roGluArgThrleuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
   |:::::::::::::::::: 111 11111111111111111111
245 CCGAAAAAACTATCGGGCTAGCGGCGCGCGAGGGAGGAATATTCACACG 294

```

```

51 GlysSerLuarGpheaAspLeuValProAsnTyrAsnProAspIleI1 67
295 AGGTATACCAATTTAAAGATCTTGCGACCTTGACCAACAATGTCGT 344
67 ePheLysAspGluLuarSergLysAlaAspArgLeuMetThrGluArgC 84
345 GTTCCGACAGAGAGACAGCGGCGCGACAGATATGACAAAGGAT 394
84 ySlySGuarGValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
395 GCCCGGATCGCTCTAAAACCTTGCGCTGATGCTGAATTAATGGCCA 444
101 GLyValArgLeuArgValThrGluGlyTrpAspGluAspGlyHisHSA1 117
445 AAAGTTAA.CTTCTTTTATCCAGCTTGACAAAAACGGAAATCCACGG 493
117 agLInAspSer.LeuHisTyrGluGlyArgValAlaLeuAspIleThrSer 133
494 T...CCACCCCTGTCATTACGAGGTAGAACAGTGCACCTCCCGGACA 540
134 AspArgAspArg.AsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluA 150
541 ACGAACCAAGCAAAAT...TTCCGAATGCTGGCGCTCCGATGGCCCTAATG 587
150 laGlyPheaAspTrpValTyrTyrGluSerArg 160
588 CCCGGTTGACTGGCTTTTACCAATCCAA 619
seq_name: gb_est14:AF201255

```

```

seq_documentation_block: 625 bp mRNA EST 29-SEP-2000
LOCUS AF201255 Botryllus schlosseri EST Botryllus schlosseri cDNA, mRNA
DEFINITION AF201255 Botryllus schlosseri EST Botryllus schlosseri cDNA, mRNA
sequence.
ACCESSION AF201255.1 GI:10419686
VERSION AF201255.1 GI:10419686
KEYWORDS EST.
SOURCE Botryllus schlosseri.
ORGANISM Botryllus schlosseri.
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
Stolidobranchia; Botryllidae; Botryllus.
REFERENCE 1 (bases 1 to 625)
AUTHORS Meyer,K.A., Leonard,P.M., Rosenberg,G.H., Miller,R.D. and Loker
,E.S.
TITLE Characterization of ESTs from the Colonial Tunicate, Botryllus
schlosseri
JOURNAL Unpublished (2000)
COMMENT Contact: Loker ES
Parasitology
Biology, University of New Mexico
Casteretter Hall, Albuquerque, NM 87131, USA
Tel: 505 277 5508
Fax: 505 277 0304
Email: esloker@unm.edu.
FEATURES
source 1..625
location/Qualifiers
/organism="Botryllus schlosseri"
/db_xref="taxon:30301"
/clone_lib="Botryllus schlosseri EST"
BASE COUNT 174 a 163 c 164 g 124 t
ORIGIN

```

```

alignment_scores:
Quality: 313.50 Length: 162
Ratio: 2.549 Gaps: 5
Percent Similarity: 75.926 Percent Identity: 41.975
alignment_block:
US-08-900-220c-17_COPY_20_198 x AF201255 ..
Align seg 1/1 to: AF201255 from: 1 to: 625

```

```

1 AlagInsercySGlyProGlyArgGlyProValGlyArgArgTyrAl 17
148 GCGGAAAGCTGC...CGGGCCCGGGCGCGCGGTTTTCACCAACACCC 194
17 aArgGlyGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
195 TCACAAACCCAAAGTGGCGGTGTATACCGGACGATATTCAAACATGT 244
34 roGluArgThrLeuGlyLaserGlyProAlaGluGlyArgValAlaArg 50
245 CCGAAAAACATATCGGGCTAGCGGCGCGCGGAGGGAATTCACAAGC 294
51 GlysSerLuarGpheaAspLeuValProAsnTyrAsnProAspIleI1 67
295 AGGTATACCAATTTAAAGATCTTGCGACCTTGACCAACAATGTCGT 344
67 ePheLysAspGluLuarSergLysAlaAspArgLeuMetThrGluArgC 84
345 GTTCCGACAGAGAGACAGCGGCGCGACAGATATGACAAAGGAT 394
84 ySlySGuarGValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
395 GCCCGGATCGCTCTAAAACCTTGCGCTGATGCTGAATTAATGGCCA 444
101 GLyValArgLeuArgValThrGluGlyTrpAspGluAspGlyHisHSA1 117
445 AAAGTTAA.CTTCTTTTATCCAGCTTGACAAAAACGGAAATCCACGG 493
117 agLInAspSer.LeuHisTyrGluGlyArgValAlaLeuAspIleThrSe 133
494 T...CCACCCCTGTCATTACGAGGTAGAACAGTGCACCTCCCGGACA 537
133 rAspArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluA 150
538 ACAACGAACCAAGCAAAATCTCCGAATGCTGGCGCTCCGATGGCCCTAATG 587
150 laGlyPheaAspTrpValTyrTyrGluSerArg 160
588 CCCGGTTGACTGGCTTTTACCAATCCAA 619
seq_name: gb_gss31:CNS04F35

```

```

seq_documentation_block: 926 bp DNA GSS 21-MAY-2000
LOCUS CNS04F35 Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 105F04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL287834.1 GI:8026342
VERSION AL287834.1 GI:8026342
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha;
Holacanthopterygii; Acanthopterygii; Percormorpha;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 926)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 926)
AUTHORS Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Winkler,P., Brotier,P., Queller,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 926)
AUTHORS Genoscope.

```

TITLE Direct Submission
JOURNAL Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
1. 926
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="105F04"
/clone_1lb="g"
/note="Genoscope sequence ID : COBGI05DC02Lp1-end : T7"

BASE COUNT 214 a 236 c 232 g 242 t 2 others
ORIGIN

alignment_scores:
Quality: 308.00 Length: 77
Ratio: 4.597 Gaps: 0
Percent Similarity: 87.013 Percent Identity: 71.429

alignment_block:
US-08-900-220C-17_COPY_20_198 x CNS04F35/rev ..

Align seg 1/1 to reverse of: CNS04F35 from: 1 to: 926

```

103 ArgLeuAlValThrGluGlyTyrAspGluAspGlyHisAlaGlnAs 119
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
926 AACCTGGCGGTCACCGAGCGCTGGAGCAGACACACACTTGGAGGA 877
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
119 pSerLeuHisTyrGluGlyArgAlaLeuAspLleThrSerAspArg 136
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
876 GTCTCTGCATACGAGCGCGCGCTGGACATCACACTCGGACCGGG 827
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
136 SPARGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGly 152
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
826 ACAAGAGCAAAATACGCGACCTGTCAGCGCTGCGAGCGCGCTTC 777
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLys 169
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
776 CACTGGCTATATATAGTATCCAAAGCTCACAATCCACTCGCTCGTAA 727
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
   ::::::::::::::::::::
726 AGGTACGAGGCGCGGAGGCGGCGGGGG 696

```

seq_name: gb_gss32:CNS05M5G

seq_documentation_block:
LOCUS CNS05M5G 819 bp DNA GSS 26-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
029G06 of library A from Tetraodon nigroviridis, genomic survey
sequence.

ACCESSION AL343645
VERSION AL343645.1 GI:8237415
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha;
Holacanthopterygii; Acanthopterygii; Percomorpha;
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 819)
AUTHORS Roest-Crolius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 819)
AUTHORS Roest-Crolius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,

TITLE Direct Submission
JOURNAL Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES
source
1. 819
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="029G06"
/clone_1lb="a"
/note="Genoscope sequence ID : COA029BD03C1-end : T7"

BASE COUNT 132 a 239 c 243 g 151 t 54 others
ORIGIN

alignment_scores:
Quality: 222.50 Length: 82
Ratio: 3.423 Gaps: 2
Percent Similarity: 79.268 Percent Identity: 62.195

alignment_block:
US-08-900-220C-17_COPY_20_198 x CNS05M5G/rev ..

Align seg 1/1 to reverse of: CNS05M5G from: 1 to: 819

```

2 GlnSerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaAr 18
   ::::::::::::::::::::::::::::::::::::::::::::::::::::
402 GAGGCGCTCGGGCGCGGGCTCGAG..GCTACGGGAAGCGGGTTCCAGCG 354
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
18 GlysGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValPro 35
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
353 GAAGTGTGCTC...CGCTCGCTTACAGAGNATTAAGCCCAAGTGGCGG 307
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
35 LuArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGly 51
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
306 AGAAGACTCTGGGGCGGCGAGATACGAAGAAATTAACGGGCAN 257
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
52 SerGluArg.PheArgAspLeuValProAspTyrAsnProAspLleLe 68
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
256 TCGAGCGCTTNAAGAGAGCTCACCCCACTACAAACCCCGACATCATCT 207
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
68 helYAspGluGluAsnSerGlyAlaAspArgLeuMetThrGlu 82
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::
206 TCAAGGATGAGGAGAAACAGGTGCGACCGCTGATGACGAG 163

```

seq_name: gb_est48:AW603741

seq_documentation_block:
LOCUS AW603741 291 bp mRNA EST 23-MAR-2000
DEFINITION CM2-CN0044-190100-073-f10 CN0044 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW603741
VERSION AW603741.1 GI:7308482
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 291)
AUTHORS HCGP <http://www.ludwig.org.br/OESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/getnum12.pl?cl=CM2&cl2=CM2-CN0044-190100-073-f10&cl3=2000-01-19&cl4=1)

Seq primer: puc 18 forward
High quality sequence start: 27
High quality sequence stop: 173.

FEATURES

Location/Qualifiers

1..291

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1ib="CN0044"

/dev_stage="Adult"

/note="Organ: colon.normal; Vector: puc18; site_1: Sma1; site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT

45 a 93 c 96 g 57 t

ORIGIN

alignment_scores:

Quality: 195.00 Length: 60
Ratio: 3.824 Gaps: 0
Percent Similarity: 85.000 Percent Identity: 61.667

alignment_block:

US-08-900-220C-17_COPY_20_198 x AM603741 ..

Align seg 1/1 to: AM603741 from: 1 to: 291

```

120 SerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgAs 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12 TCCTCGCTTATGACGCGCGCGGTGACATCACACATCACGACCGCGA 61
136 PARGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPheA 153
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 CCGCATACGATGACTGCTGGCGCGCTTGCGACGCGCGCGCTTTG 111
153 sPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAla 169
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 ACTGGTGTCTTACGACAAAGGCCCGCATGTATGCTCCGTCTGCTCC 161
170 AspAsnSerLeuAlaValArgAlaGlyGly 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GAGCAGCTGCGCGCGACGACCAAGACGCGCGCG 191

```

seq_name: gb_est15:AI015155

seq_documentation_block:

LOCUS AI015155 455 bp mRNA EST 27-AUG-1998
DEFINITION O70h05.s1 Soares-total_fetus.Nb2HF8_9w Homo sapiens cDNA clone
PROTEIN:1622169 3' similar to TR:Q98938 Q98938 INDIAN HDGEHOG
PROTEIN:,, mRNA sequence.

ACCESSION AI015155
VERSION AI015155.1 GI:3229491

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 455)
REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

Insert Length: 615 Std Error: 0.00
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 392.

FEATURES

Location/Qualifiers

1..459

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_1ib="Soares-total_fetus_Nb2HF8_9w"

/dev_stage="8-9 weeks"

/lab_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from pooled 8-9 week (total) fetus material with a Not I - Oligo(dT) primer (5' TGTTCATCTGAGTGGAGCGCGCGCTTATGCTCCGTCTGCTCC 3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia). digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

106 a 127 c 133 g 93 t

ORIGIN

alignment_scores:

Quality: 180.00 Length: 61
Ratio: 3.673 Gaps: 3
Percent Similarity: 80.328 Percent Identity: 67.213

alignment_block:

US-08-900-220C-17_COPY_20_198 x AI015155/rev ..

Align seg 1/1 to reverse of: AI015155 from: 1 to: 459

```

5 GlyProGlyArgGlyProValGlyArgArg.ArgTyrAlaArgLysInt 21
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459 GGGCGGGGTGG...GTGGTGGGCGACCGCGCGCGCGCGCGCGCAA...C 416
21 euValProLeuLeuTyrLysGlnPheValProGlyValProGlyArgThr 37
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 TCGGCGCGCTGCGCTACAGACAGTTCAGCCCAATGTGCGCGAAGACC 366
38 LeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgLysSerGluAr 54
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 CTGGCGCGCGCGACGCGCGCTATGAAGGCAAGATCGCTGCGACGCGAGCG 316
54 gpheArgAspLeuValProAsnTyrAsnPro 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 CTTCATGAGCTCACCCCGCAATTACATCCA 285

```

seq_name: gb_est38:AV604922

seq_documentation_block:

LOCUS AV604922 555 bp mRNA EST 30-AUG-2000
DEFINITION AV604922 Bos taurus kidney fetus Bos taurus cDNA clone EIKI025F10
5', mRNA sequence.

ACCESSION AV604922
VERSION AV604922.1 GI:9735295

KEYWORDS

EST.

SOURCE

cow.

ORGANISM

Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 555)
REFERENCE Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitohzono, A. and

TITLE Suzuki, H.
JOURNAL bovine CDNA sequencing
COMMENT Unpublished (2000)
Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
OdaKura, Nishigo, Nishi-Shirakawa, Fukushima 961-8061, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.

FEATURES This clone was obtained from a polyA-deleted cDNA library.

source Location/Qualifiers
1.355

/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="E1K1025P10"
/clone_lib="Bos taurus kidney fetus"
/tissue_type="kidney"
/dev_stage="fetus"
/lab_host="DH10B"
/note="Vector: pZ1; Site_1: SalI; Site_2: NotI; Poly A was deleted from a NotI site"
BASE COUNT 98 a 184 c 175 g 98 t
ORIGIN

alignment_scores:

Quality: 165.00 Length: 46
Ratio: 4.024 Gaps: 1
Percent Similarity: 89.130 Percent Identity: 73.913

alignment_block:

US-08-900-220C-17_COPY_20_198 x AV604922 ..

Align seg 1/1 to: AV604922 from: 1 to: 555

```

135 ArgAspArgAnlySTyrgYLeuLeuAla..ArgLeuAlaValGluAla 150
|||||
10 CGCGACCGCAGTAAGTACGACGACGACGCGGTTGCGAGTGGAGCGCC 59
|||||
151 GLYpHeAsPTpValTYrTYrGlusErArGAsnHsVAlHsVAlSerVa 167
|||||
60 GCGTTCAGCTGGGTATTAAGAGTCAAGGCCAGCGTCTGCTCGT 109
|||||
167 LlysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||
110 CAACTCCGACGACCTAGCGCCGACCAAGACAGTGGC 146

```

seq_name: gb_est54:BB030536

seq_documentation_block:

LOCUS BB030536 239 bp mRNA EST 23-JUN-2000
DEFINITION BB030536 RIKEN full-length enriched, adult male thymus Mus musculus
cDNA clone 5830455115 3' similar to AF148226 Rattus norvegicus
desert hedgehog protein (dhn) mRNA, mRNA sequence.

ACCESSION BB030536
VERSION BB030536.1 GI:8385287

KEYWORDS:

EST.
house mouse.
Mus musculus

SOURCE

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 239)

REFERENCE

AUTHORS Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, H., Hayatsu, N.,
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomiwa, N., Toya

TITLE
JOURNAL
COMMENT

T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,
M., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Kono, H., et al.)
Unpublished (2000)
Contact: Yoshinori Hayashizaki
Genome Exploration Research Group, Life Science Tsukuba Center,
Genome Science Laboratory
The Institute of Physical and Chemical Research (RIKEN), Genomic
Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: +81-298-36-9013
Fax: +81-298-36-9038
Email: genome-res@rtc.riken.go.jp,
URL: http://genome.rtc.riken.go.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunari, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,
Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

source

Location/Qualifiers
1.239

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5830455115"
/clone_lib="RIKEN full-length enriched, adult male thymus"
/sex="male"
/tissue_type="thymus"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTTAATTAATTAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified Bluescript KS(+) after bulk excision from lambdaBda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT 60 a 72 c 48 g 59 t
ORIGIN

alignment_scores:

Quality: 160.00 Length: 39
Ratio: 4.444 Gaps: 0
Percent Similarity: 92.308 Percent Identity: 71.795

alignment_block:

US-08-900-220C-17_COPY_20_198 x BB030536/rev ..

Align seg 1/1 to reverse of: BB030536 from: 1 to: 239

```

83 ArgCysIysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTr 99
|||||
138 CGTTGCAAGAGAGGGGTGAATGCTTAATCCATCGCGGTGATAAATATATGTG 89
99 pProGlyValArgLeuArgValThrGluGlyTTPaspGluAspGlyHisH 116
| |||||
88 GCCCGGAGTACGGGTAGTGTGAATGAAGGGGTGGGAAAAAGATGGGCACC 39
116 tAlaGlnAspSerLeu 121
|||||
38 ATGCTTAGGAACCTCTC 22

```

OM of: US-08-900-220c-17_COPY_20_198 to: N.Geneseq_0401:* out_format : pfs
Date: Apr 22, 2001 12:07 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+pn.model -DEV=xlp
-O=/cgn2.1/uspt0.spool/US08900220/rnat.19042001.104047.21703/app-query.fasta_1.253
-DB=N.Geneseq_0401 -QFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOCPCL=0.000 -ICOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FCGAPOP=6.000 -FCGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORMEXT -MINLEN=0 -MAXLEN=200000000
-USER=US08900220_cgn1_1186 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-08-900-220c-17_COPY_20_198
Query length: 179
Database: N.Geneseq_0401:*
Database sequences: 678276
Database length: 291890651
Search time (sec): 111.600000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	! Documentation	..	
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25624				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25105				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07270				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X50453				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A27883				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A20281				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X52264				950.00	1874.66	1.8e-96	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X00317				950.00	1870.38	3.1e-96	1820
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A58931				950.00	1870.38	3.1e-96	1820
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X62396				949.00	1880.49	8.3e-97	548
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X62397				949.00	1879.54	9.4e-97	602
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X62395				949.00	1872.52	2.3e-96	1188
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X62393				946.00	1854.92	2.2e-95	528
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X62394				936.00	1847.32	5.9e-95	1122
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X01642				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25618				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07272				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07273				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X16183				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A50447				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A27877				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A30275				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X52598				932.00	1838.75	1.8e-94	1190
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25099				932.00	1838.74	1.8e-94	1191
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25617				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25098				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07271				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X16182				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A50446				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A27876				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A30274				741.00	1456.97	3.2e-73	1277
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X525257				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25623				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25104				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07269				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X16188				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A50452				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A27882				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A30280				740.50	1453.56	5.0e-73	1622
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:X25263				740.50	1453.56	5.0e-73	1622

/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X01637 + 740.00 1454.69 4.3e-73 1313
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25101 + 740.00 1454.69 4.3e-73 1313
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X07274 + 740.00 1454.69 4.3e-73 1313
/SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X16185 + 740.00 1454.69 4.3e-73 1313
/SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT:A50449 + 740.00 1454.69 4.3e-73 1313

seq_name: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT:X25624

seq_documentation_block:

ID X25624 standard; cDNA; 1190 BP.

AC X25624;

DT 02-AUG-1999 (first entry)

DE Human Desert hedgehog Dhh protein cDNA.

XX Desert hedgehog; Dhh; human; epithelial tissue; epithelium;

RW cutaneous tissue; skin; hair; wound healing; vulnertax;

RW burn; skin grafting; pressure sore; ulcer; ulcerative colitis;

RW alopecia; psoriasis; keratosis; acne; comedogenic lesion;

RW folliculitis; pseudofolliculitis; keratoacanthoma; callosities;

RW Darier's disease; scar; autoimmune disease; pemphigus;

RW epidermolysis; lupus lesion; desquamative lesion; carcinoma;

RW therapy; hedgehog therapeutic; plc therapeutic; patched; ds.

XX Homo sapiens.

XX W09920298-A1.

PD 20-OCT-1998; 98WO-US22227.

XX 11-SEP-1998; 98US-0151999.

PR 20-OCT-1997; 97US-0935352.

XX (ONTO-) ONTOGENY INC.

XX Wang EA;

DR WPI: 1999-288170/24.

XX P-PSDB; Y05861.

XX Use of hedgehog polypeptides on patched therapeutics

PS Claim 29; Page 115-117; 146pp; English.

XX This nucleotide sequence comprises a coding region for human

Desert hedgehog protein Dhh (see Y05861). The invention relates to

a method for modulating the growth state of an epithelial cell by

ecologically contacting the epithelial cell, in vitro or in vivo,

with a hedgehog therapeutic (i.e. a hedgehog polypeptide or gene

therapy construct) or plc therapeutic (i.e. a small organic molecule

that mimics the effect of hedgehog proteins on patched signalling, or

activates or potentiates patched signalling) in an amount effective

to alter the rate of proliferation of the epithelial cell. The

hedgehog therapeutic preferably comprises at least a bioactive

extracellular portion of a hedgehog protein (see Y05861-62) encoded

by a vertebrate hedgehog gene (see X25617-25), especially a human

hedgehog gene. Promotion of proliferation of epithelial cells can

be used to control a wound healing process in e.g. burn treatment,

skin regeneration, skin grafting, pressure sore treatment, dermal

ulcer treatment, post surgery scar reduction or treatment of

ulcerative colitis (claimed). It can also be used to induce hair

growth for the treatment of alopecia (claimed). Inhibition of the

growth of epithelial tissue can be used to treat or prevent

hyperplastic or neoplastic conditions, e.g. psoriasis, keratosis,

acne, comedogenic lesions, folliculitis and pseudofolliculitis,

keratoacanthoma, callosities, Darier's disease, keloids,

hypertrophic scars, or autoimmune disorders, e.g. aphthous ulcers,

pemphigus vulgaris, pemphigus foliaceus, pemphigus vegetans,

pemphigus erythematous, epidermolysis, lupus lesions, desquamative

CC lesions or carcinomas. The methods can also be used to counteract
CC the effects of ageing on skin.

XX Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment_scores:

Quality: 950.00 Length: 179
Ratio: 5.307 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-900-220C-17_COPY_20_198 x X25624 ..

Align seg 1/1 to: X25624 from: 1 to: 1190

```

1  ALaGInSerCySGlyProGlyAArgGlyProValGlyAArgAArgTYrAl 17
   |||||||
58  GCGCAGAGCTCGGGCGCGGGCGGGCGGTGGCGCGCGCGCTATGC 107
   |||||||
17  AArgLySGInLeuValProLeuLeuTYrLySGInPheValProGlyValP 34
   |||||||
108  GCGCAGAGAGCTCGTCCGCTACTCTACACAAATTGTGCGCGCGCTGC 157
   |||||||
34  roGluArgThLeuGlyAlaSerGlyProAlaGluGlyAArgValAlaArg 50
   |||||||
158  CAGAGCGGACCTGGGCGCCAGTGGCGGAGGAGGAGGAGGTGGCAAGG 207
   |||||||
51  GlySerGluArgPheArgAspLeuValProAsnTYrAsnProAspIleI 67
   |||||||
208  GCGTCGCGAGCGCTTCCGGGACCTCGTGCCCAACATACACCCCGACATCAT 257
   |||||||
67  ePheLyAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
   |||||||
258  CTTCAAGAGTGAAGAGAACACTGGAGCGGACCGCTGATGACCGAGCTTT 307
   |||||||
84  ySLySGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
   |||||||
308  GCAAGGAGAGGCTGACCGCTTTGGCCATTGGCTGATGATGATGATGGCCC 357
   |||||||
101  GLyValArgLeuArgValThrGluGlyTrpAspGluAspGlyHisHisAl 117
   |||||||
358  GGAGTCGCGCTACAGATCACTAGAGGCTGGGACAGAGCGGCCACACAGC 407
   |||||||
117  aGlnAspSerLeuHisTYrGluGlyArgAlaLeuAspIleThrThrSerA 134
   |||||||
408  TCAGGATTCACCTCCACTACGAAAGCCGCTTTGGACATCACTACGCTCG 457
   |||||||
134  sPArgAspArgAsnLySTyGlyLeuLeuAlaArgLeuAlaValGluAla 150
   |||||||
458  ACCCGGAGCGGACACAGATATGGTTGCTGGCGCGCTCGCAGTGGAGACC 507
   |||||||
151  GLyPheAspTrpValTYrTYrGluSerArgAsnHisValHisValSerVa 167
   |||||||
508  GCGTTTCGACTGGGTCTACTACGAGTCCCGCAACACAGCTCCACGTTGGGT 557
   |||||||
167  lLyAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
   |||||||
558  CAAAGCTGATACCTACCTGCGGCGCGCGGCGGCGG 594

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.X25105

seq_documentation_block:

ID X25105 standard; cDNA: 1190 BP.
AC X25105;
XX
DT 05-JUL-1999 (first entry)
XX
DE Human Desert hedgehog protein Dhh cDNA.
XX
KW Desert'hedgehog; Dhh gene; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;

KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.

XX Homo sapiens.

XX MO9910004-A2.

XX 04-MAR-1999.

XX 28-AUG-1998; 98MO-US17922.

XX 29-AUG-1997; 97US-0057394.

XX (ONTO-) ONTOGENY INC.

XX Bladgen CS, Currie PD, Hughes SM, Ingham FW;

XX WPI: 1999-243557/20.

XX P-PsDB; Y05517.

XX A new method to regulate muscle growth

XX Disclosure: Page 106-108; 130pp; English.

This nucleotide sequence comprises a coding region for the human
Desert hedgehog protein Dhh (see Y05517). The invention relates to
a method for modulating the formation and/or maintenance of muscle
tissue by ecotopically contacting muscle cells, especially muscle
stem/progenitor cells, in vitro or in vivo, with a hedgehog
therapeutic (i.e. hedgehog polypeptides and gene therapy
constructs) or ptc therapeutic (i.e. a small organic molecule that
mimics the effect of hedgehog proteins on patched signalling, or
activates or potentiates patched signalling) in an amount effective
to alter the growth state of the treated cells. Also claimed is a
method for treatment or prevention of disorders of, or surgical or
cosmetic repair of, such muscle tissues, by administering a
hedgehog polypeptide or ptc therapeutic. The disorder may be
muscle atrophy, in particular skeletal muscle atrophy or cardiac
muscle atrophy, cachexia, or muscular myopathy (all claimed). The
hedgehog polypeptide or ptc therapeutic can inhibit growth of
myoblastic-derived tissue to provide treatment of hyperplastic or
neoplastic growth of muscle tissue such as in myoblastic sarcoma
CC (also claimed). The hedgehog therapeutic preferably comprises at
least a bioactive extracellular portion of a hedgehog protein (see
Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),
especially a human hedgehog gene.

Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment_scores:

Quality: 950.00 Length: 179
Ratio: 5.307 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-900-220C-17_COPY_20_198 x X25105 ..

Align seg 1/1 to: X25105 from: 1 to: 1190

```

1  ALaGInSerCySGlyProGlyAArgGlyProValGlyAArgAArgTYrAl 17
   |||||||
58  GCGCAGAGCTCGGGCGCGGGCGGGCGGTGGCGCGCGCGCTATGC 107
   |||||||
17  AArgLySGInLeuValProLeuLeuTYrLySGInPheValProGlyValP 34
   |||||||
108  GCGCAGAGAGCTCGTCCGCTACTCTACACAAATTGTGCGCGCGCTGC 157
   |||||||
34  roGluArgThLeuGlyAlaSerGlyProAlaGluGlyAArgValAlaArg 50
   |||||||
158  CAGAGCGGACCTGGGCGCCAGTGGCGGAGGAGGAGGTGGCAAGG 207
   |||||||
51  GlySerGluArgPheArgAspLeuValProAsnTYrAsnProAspIleI 67
   |||||||

```



```

208 GCCTCCGAGCGCTTCGGGACCCTCGTCCCCAATCAACCCCGACATCAT 257
67 ephelYsAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAAGATGAGGAGAACAGTGAAGCGGACCGCTGATGACGAGCGCT 307
84 yslYsGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGGAGAGGGTGAACGCTTGGCCATTCGCTGATGAACTATGGGCC 357
101 GlyValArgLeuArgValThrGluGluTrpAspGluAspGlyHisAla 117
|||||
358 GGAGTGGCCCTACGAGTACGAGTGAAGGCTGGGACGAGGAGCCGACACCC 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSerA 134
|||||
408 TCAGATTCACATCCACTAGCAAGGCGGTGCTTGACATCATAGTGTG 457
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGCGACCGGACAAAGTATGGTGTGCGCGCCGCTGCAAGTGAAGCC 507
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVa 167
|||||
508 GGCTTCGACGTGGGTCTACTACGAGTCCGCAACACGCTCCACGTGTGGT 557
167 lLysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||
558 CAAGCTGATTAATCACTGCGGCGGTCCGGGCGGCGGCGG 594

```

seq.name: /SIDS2/gcdata/geneseq/geneseqn/NA1999.DAT:X07270

seq_documentation_block:
ID X07270 standard; cDNA: 1190 BP.

```

XX AC X07270;
XX DT 21-MAY-1999 (first entry)
XX DE Human Desert hedgehog (Dhh) cDNA.
XX KW Desert hedgehog; Dhh gene; human; dopaminergic; GABA-nergic;
KW ptc therapeutic; patched; signal transduction; Parkinson's disease;
KW Huntington's disease; amyotrophic lateral sclerosis;
KW cerebral ischaemia; hypoxia; neuroprotective; gene therapy; ss; ds.
XX OS Homo sapiens.
XX PN W09904775-A2.
XX PD 04-FEB-1999.
XX PF 24-JUL-1998; 98MO-US15419.
XX PR 24-JUL-1997; 97US-0900220.
XX PA (ONTO-) ONTOGENY INC.
XX PI Mahanthappa NK, Miao N, Pang K, Wang M:
XX DR WPI, 1999-142578/12.
XX DR P-PSDB; W97764.
XX PT Increasing the survival of neuronal, dopaminergic and GABA-nergic
XX PT cells - by using a ptc therapeutic such as a protein kinase
XX PT inhibitor, or an agent derived from hedgehog polypeptides, useful in
XX PT the treatment of Parkinson's disease
XX PS Disclosure; Page 97-99; 138pp; English.
XX CC This nucleotide sequence comprises a coding region for the human
XX CC Dhh Desert hedgehog protein (see W97764). The invention is based on
XX CC the finding that hedgehog proteins are useful as protective agents

```

CC in the treatment and prophylaxis of neurodegenerative disorders
CC resulting from the loss of dopaminergic and/or GABA-nergic neurons,
CC or the general loss of tissue from the substantia nigra.
CC Exemplary disorders include Parkinson's disease, Huntington's
CC disease (both claimed), amyotrophic lateral sclerosis and cerebral
CC ischaemia. The invention relates to hedgehog therapeutics (i.e.
CC hedgehog polypeptides and gene therapy constructs e.g. constructs
CC encoding recombinant hedgehog polypeptides and trans-activation
CC constructs for altering hedgehog gene regulatory sequences) and
CC ptc therapeutics (i.e. agents which mimic the effect of naturally
CC occurring hedgehog proteins on patched signalling) that are
CC effective in both human and animal subjects. A bioactive
CC polypeptide comprising amino acid residues 23-198 of human Dhh is
CC preferred. The products can also be used for the maintenance of
CC differentiated neurons in cultures, and to enhance the implantation
CC of such neuronal cells in an animal. They can also be used to
CC prevent or treat neurodegenerative conditions arising from the use
CC of certain drugs, and in the prevention and/or treatment of hypoxia,
CC e.g. as a neuroprotective agent.

XX Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment_scores:
Quality: 950.00 Length: 179
Ratio: 5.307 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-900-220c-17_COPY_20_198 x X07270 ..

Align seg 1/1 to: X07270 from: 1 to: 1190

```

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
|||||
58 GCCCAGAGCTGGCGGGCGGCGGGCGGGTGGCGGCGCGCTATGTC 107
17 aArgGlyGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
|||||
108 GCGCAAGCAGCTGCTGCCGCTACTACAAAGCAATTGTGCCGGCGGTGC 157
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CAGAGCGGAGCCCTGGCGGCGGCGGCGGCGGAGGAGGCGGCGCAAG 207
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
208 GGCTCCGAGCGCTTCGGGACCTGTGCCCCAATCAACCCCGACATCAT 257
67 ephelYsAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAAGATGAGGAGAACAGTGAAGCGGACCGCTGATGACGAGCGCT 307
84 yslYsGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGGAGAGGGTGAACGCTTGGCCATTCGCTGATGAACTATGGGCC 357
101 GlyValArgLeuArgValThrGluGluTrpAspGluAspGlyHisAla 117
|||||
358 GGAGTGGCCCTACGAGTACGAGTGAAGGCTGGGACGAGGAGCCGACACCC 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSerA 134
|||||
408 TCAGATTCACATCCACTAGCAAGGCGGTGCTTGACATCATAGTGTG 457
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGCGACCGGACAAAGTATGGTGTGCGCGCCGCTGCAAGTGAAGCC 507
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVa 167
|||||
508 GGCTTCGACGTGGGTCTACTACGAGTCCGCAACACGCTCCACGTGTGGT 557

```

167 LysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
 558 CAAAGCTGATACCTACTGCGGCGGCGGCGGC 594

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:A50453

seq_documentation_block:

ID A50453 standard; cDNA: 1190 BP.

AC A50453:

DT 05-DEC-2000 (first entry)

DE Human Desert hedgehog 1th cDNA.

Desert hedgehog; Dhh; human; agonist; antagonist;
 lipid modulator; vacuole; cholesterol disorder; lipid disorder;
 lipid metabolism; lipid storage; lipid transport; apolipoprotein;
 triglyceride; hypercholesterolemia; abetalipoproteinemia;
 hypobetalipoproteinemia; chylomicron retention; Anderson's disease;
 fat absorption; atherosclerosis; obesity; weight loss;
 vitamin A disorder; vitamin E disorder; anlipemia; anorectic;
 antiarteriosclerotic; gene therapy; diagnosis; ss.

OS Homo sapiens.

PN WO200051628-A2.

PD 08-SEP-2000.

PF 03-MAR-2000; 2000MO-US05662.

PR 03-MAR-1999; 990S-0122640.

PR 15-MAR-1999; 990S-0124446.

PA (BIOJ) BIOGEN INC.

PI Burkly L, Wang LC;

DR WPI: 2000-611340/58.

DR P-PSDB: Y95979.

Use of lipid modulators (e.g. hedgehog agonists or antagonists) for
 modulating lipid metabolism and storage, especially useful for treating
 lipid metabolism or cholesterol disorders, e.g. obesity or
 hypercholesterolemia -

Disclosure: Page 102-104; 136pp; English.

The present sequence of that of cDNA coding for human Desert
 hedgehog (Dhh) protein (see Y95979). The invention provides
 claimed methods for modulating lipid metabolism, for modulating
 vacuole formation in intestinal epithelial cells, for modulating
 the accumulation of fat in intestinal epithelial cells, for
 treating a cholesterol disorder and for treating a lipid metabolism
 disorder in an animal (especially a human) by administering a lipid
 modulator selected from a hedgehog antagonist or hedgehog agonist.
 In particular, the lipid metabolism disorder is a lipid storage
 disorder, a lipid transport disorder, an apolipoprotein disorder,
 a triglyceride disorder, e.g. a triglyceride metabolism disorder, a
 triglyceride transport disorder or a triglyceride storage disorder;
 a diet-induced hypercholesterolemia, hypercholesterolemia,
 abetalipoproteinemia, hypobetalipoproteinemia; a
 chylomicron-retention disorder; Anderson's disease, a fat
 absorption disorder, e.g. obesity or associated with weight loss,
 normotriglyceridemic abetalipoproteinemia, an apolipoprotein-B100
 deficiency, a fat soluble vitamin disorder, where the fat soluble
 vitamin is vitamin A or E, or atherosclerosis (all claimed). The
 hedgehog antagonist binds to the hedgehog receptor, but does not
 elicit a response. It is preferably a hedgehog mimetic, a modified
 hedgehog protein or an anti-hedgehog homologue, especially a human
 antibody, chimeric antibody or humanised antibody. The methods are
 useful in preventing these disorders or protecting a subject from

these disorders. The hedgehog antagonist and agonist are also
 useful in diagnosis and research associated with these disorders.
 The lipid modulators may also be used as a part of a gene therapy
 protocol to deliver polynucleotides encoding these lipid
 CC modulators. Hedgehog gene constructs can also be used for
 CC recombinant production of hedgehog polypeptides.

Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment_scores:

Quality:	950.00	Length:	179
Ratio:	5.307	Gaps:	0
Percent Similarity:	100.000	Percent Identity:	100.000

alignment_block:

us-08-900-220c-17_copy_20_198 x A50453 ..

Align seq 1/1 to: A50453 from: 1 to: 1190

1 AlaglnSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
 |||||
 58 GCCACAGAGCTGCGGGCGCGGGCGGGCGGGCTGGCGCGCGCTATGC 107
 17 aArglyGlnLeuValProLeuLeuTyrlyGlnPheValProGlyValP 34
 |||||
 108 GCGCAGACCTGCTGCGGCTACTCTCAAGCAATTTGTGCGCGCGCTGC 157
 34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
 |||||
 158 CAGAGCGGACCTGCGGGCGCGGACGAGCGGAGGAGGAGGAGGAGG 207
 51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle 67
 |||||
 208 GCGTCGAGCGCTTCCGGACCTGCTGCCCACTACCAACCCGACATCAT 257
 67 ePheLysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
 |||||
 258 CTTCAGAGATGAGAGAAAGTGGAGCGGACCGCTGTATACCGAGGCTT 307
 84 ySLysGluArgValAlaAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
 |||||
 308 GCAAGGAGAGGTGACACGCTTGGCCATTGGCCGATGAAACATGTGCCC 357
 101 GlyValArgLeuArgValThrGluGlyTyrPaspGluAspGlnIshIshAl 117
 |||||
 358 GGAATGCGCTACGAGTGAAGGAGGCTGGGAGGAGGAGGAGGAGGAG 407
 117 acGlnAspSerLeuHisTyrGluArgAlaAlaLeuAspIleThrThrSera 134
 |||||
 408 TCAGGATTCACCTCCACTAGCAAGGCGGTGCTTGGACATCACTACGCTG 457
 134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValAlaGluAla 150
 |||||
 458 ACCGCGACCCCAACACTATGAGGTGCTGGCCGCCCTCGGAGTGAAGCC 507
 508 GCGTTTCACCTGAGGTCTACTACGAGTCCCGCAACACAGCTCCACGTGCGGT 557
 167 LysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
 |||||
 558 CAAAGCTGATACCTACTGCGGCGGCGGCGGC 594

seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:A27883

seq_documentation_block:

ID A27883 standard; cDNA: 1190 BP.

AC A27883:

DT 12-SEP-2000 (first entry)

DE Human Desert hedgehog Dhh cDNA.
 XX
 KW Desert hedgehog; Dhh; human; excitotoxicity; Parkinson's disease;
 KM Huntington's disease; neuronal degeneration; neuroprotective;
 KM dopaminergic; GABAergic; substantia nigra; gene therapy; ss.
 XX
 OS Homo sapiens.
 PN W0200035948-A1.
 XX
 PD 22-JUN-2000.
 XX
 PF 03-DEC-1999; 99WO-US28721.
 XX
 PR 03-DEC-1998; 98WO-US25676.
 PR 27-JAN-1999; 99US-0238243.
 PR 03-JUN-1999; 99US-0325602.
 XX
 PA (BIOJ) BIOGEN INC.
 PI (ONTO-) ONTOGENY INC.
 XX
 PI Galdes A, Mahanthappa N;
 DR MPI: 2000-431570/37.
 DR P-PSDB; Y95288.
 XX
 XX
 PT Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,
 PT senile dementia and Korsakoff's disease, by using lipophilic modified
 PT hedgehog polypeptide -
 XX
 PS Disclosure; Page 145-147; 174pp; English.
 XX
 CC The present sequence of that of cDNA coding for human Desert
 CC hedgehog (Dhh) protein (see Y95288). The invention relates to a
 CC method for promoting the survival and/or functional performance
 CC of neuronal cells, especially substantia nigra, dopaminergic or
 CC GABAergic neurons that are susceptible to exotoxicity, by
 CC contacting the cells, in vitro or in vivo, with a lipophilic
 CC (e.g. cholesterol) modified hedgehog polypeptide. The method is
 CC used to treat or prevent Parkinson's disease, Huntington's disease,
 CC domoic acid poisoning, spinal cord trauma, hypoglycemia, mechanical
 CC trauma to the nervous system, senile dementia, Korsakoff's disease,
 CC schizophrenia, AIDS dementia, multi-infarct dementia, mood
 CC disorders, depression, chemical toxicity, neuronal damage
 CC associated with uncontrolled seizures such as epileptic seizures,
 CC neuronal injury associated with HIV and AIDS, neurodegeneration
 CC associated with Down's syndrome, neuropathic pain syndrome,
 CC olivopontocerebral atrophy, amyotrophic lateral sclerosis,
 CC mitochondrial abnormalities, Alzheimer's disease, hepatic
 CC encephalopathy, Tourette's syndrome and drug addiction (all
 CC claimed). The lipophilic modified hedgehog polypeptide is also
 CC useful for promoting survival and/or functional performance of
 CC neuronal cells susceptible to exotoxicity. Hedgehog gene
 CC constructs can be used for recombinant production of hedgehog
 CC polypeptides, to design probes and primers, and in gene therapy
 CC protocols to deliver nucleic acids encoding a neuroprotective
 CC form of hedgehog polypeptide.
 XX
 SO Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

Alignment scores:
 Quality: 950.00 Length: 179
 Ratio: 5.307 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:
 US-08-900-220C-17_COPY_20_198 x A27883 ..
 Align seg 1/1 to: A27883 from: 1 to: 1190

1 AAGInserCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17
 ||||||||||||||||||||||||||||||||||||||||||||||||

58 GCCCAGAGCTGCGGGCGGGCGGGCGGGCTTGCGGGCGGGCGGCTATGC 107
 17 AArgLysGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
 ||||||||||||||||||||||||||||||||||||||||||||||||
 108 GCGCAAGAGCATGCTGCTCCGCTACTACTACAAAGCAATTGTGCCGGGCTCC 157
 34 roGLuArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
 ||||||||||||||||||||||||||||||||||||||||||||||||
 158 CAGAGCGGAGCCCTGGGCGGCGGAGTGGGCGGAGCGGAGGGGTGGCAAG 207
 51 GlySerGluArgPheArgAspLeuValProAsnTyrAspAspIleI 67
 ||||||||||||||||||||||||||||||||||||||||||||||||
 208 GGCTCCGAGCGCTCCGGGACCTCTGTCCTCCCACTACAAACCCGACCTTCAT 257
 67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
 ||||||||||||||||||||||||||||||||||||||||||||||||
 258 CTTCAGAGATGAGAGAACAGTGGAGCCGACCCCTGATGACCGAGCGCTT 307
 84 yslYsgLluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
 ||||||||||||||||||||||||||||||||||||||||||||||||
 308 GCAAGAGAGAGGTGAGACGCTTGCCCATTCGCGTGAACATGTGGCCC 357
 101 GlyValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAla 117
 ||||||||||||||||||||||||||||||||||||||||||||||||
 358 GGAGTGGCTGCTAGAGTGAAGTGAAGGCTGGAGAGAGAGGACCCACCC 407
 117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSera 134
 ||||||||||||||||||||||||||||||||||||||||||||||||
 408 TCAGAGATCACTCCACTACGAAGCGGCTGTTGGACATCAGTACTGTCG 457
 134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
 ||||||||||||||||||||||||||||||||||||||||||||||||
 458 ACCGCCACCGCAACAAGTATGGGTCTGCGCGCGCTCCAGAGTGAAGCC 507
 151 GlyPheAspTyrValTyrTyrGluSerArgAsnHisValHisValSera 167
 ||||||||||||||||||||||||||||||||||||||||||||||||
 508 GGCTTGACTGGGTCTACTAGAGTCCCGCAACACGTCACGTCGCTGGGT 557
 167 llySAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
 ||||||||||||||||||||||||||||||||||||||||||||||||
 558 CAAGCTGATTACTCACTGGCGGTCCGGCGGCGGCGGCG 594

seq.name: /SID52/gcdata/geneseq/geneseqn/NA2000.DAT:A30281

seq_documentation_block:
 ID A30281 standard; cDNA; 1190 BP.
 XX
 AC A30281;
 XX
 DT 11-SEP-2000 (first entry)
 XX
 DE Partial human Dhh coding sequence.

Human; desert hedgehog; Dhh; neuromuscular disorder; neuropathy;
 Guillian-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;
 Chronic inflammatory demyelinating polyneuropathy; CIPD;
 gene therapy; infection; inflammation; hereditary neuropathy;
 Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;
 multiple myeloma; nutritional imbalance; kidney disease;
 hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;
 Tangle disease; Krabbe's disease; Metachromatic leukodystrophy;
 Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;
 Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;
 Hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;
 Waidenström's Macroglobulinemia; Chronic Lymphocytic Leukemia;
 neuroprotective; cytoprotective; patched-mediated signal transduction;
 ss.
 KM
 KM
 KM
 OS Homo sapiens.
 XX
 XX
 Key Location/Qualifiers
 CDS 1..1188
 FT /*tag= a

```
FT      /partial
FT      /product= "human DhH"
XX      WO200027422-A2.
XX      18-MAY-2000.
XX      08-NOV-1999: 99WO-US26334.
XX      06-NOV-1998: 98US-0187387.
XX      (BIOJ ) BIOGEN INC.
XX      (ONTO-) ONTOGENY INC.
XX      Galdes A, Mahanthappa N:
XX      MPI: 2000-387341/73.
XX      P-PSDB: Y96250.
XX
XX      Novel method of preventing deterioration of peripheral nerves, useful
XX      for treating or preventing neuropathy, e.g. where associated with
XX      diabetes or viral infection, by administering hedgehog or patched agent
XX
XX      Claim 10: Page 123-125; 152pp; English.
XX
XX      The present sequence is the partial human coding sequence for desert
XX      hedgehog, DhH. The encoded protein inhibits expression of the patched
XX      gene which has been implicated in neuromuscular disorders (neuropathies).
XX      This sequence may therefore be used for treating neuromuscular disorders
XX      i.e. preventing degradation in function of motor or sensory nerves and
XX      protecting peripheral nerve cells under conditions that normally cause
XX      neuropathy. A variety of neuromuscular disorders may be treated:
XX      Guillain-Barre syndrome, GBS; peripheral neuropathy; diabetic neuropathy;
XX      alcohol-induced neuropathy; chronic inflammatory demyelinating
XX      polyneuropathy, CIPD; infection-induced neuropathy, including HIV
XX      infection; inflammation-induced neuropathy; hereditary neuropathy, e.g.
XX      Charcot-Marie-Tooth disease (CMT), Familial Amyloidotic neuropathy,
XX      Refsum's disease, Abetalipoproteinemia, Tangier disease, Krabbe's
XX      disease, Metachromatic Leukodystrophy, Fabry's disease, Dejerine-Sottas
XX      syndrome, Hereditary sensory neuropathy Type II (HSN II) and Amyotrophic
XX      lateral sclerosis (ALS); acute neuropathy e.g. Miller-Fisher syndrome;
XX      neuropathy caused by vasculitis; neuropathy associated with tumours e.g.
XX      lung cancer, multiple myeloma, B-cell lymphoma, Waldenstrom's
XX      Macroglobulinaemia, Chronic Lymphocytic Leukaemia; neuropathy associated
XX      with: amyloidosis, nutritional imbalance, kidney disease, trauma; and
XX      hypothyroid neuropathy. This sequence may be used in gene therapy of the
XX      above disorders.
XX
XX      Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;
XX
XX      alignment_scores:
XX          Quality: 950.00      Length: 179
XX          Ratio: 5.307      Gaps: 0
XX      Percent Similarity: 100.000      Percent Identity: 100.000
XX
XX      alignment_block:
XX      US-08-900-220C-17_COPY_20_198 x A30281 ..
XX
XX      Align seg 1/1 to: A30281 from: 1 to: 1190
XX
XX      1 AAlaInserCySGlYProGlyAArgGlyProValGAlYArGaRgArGTYrAl 17
XX      |||||||
XX      58 GCCCAGAGCTGCGCGCGCGCGCGCGCGTGGCGCGCGCGCTATGCG 107
XX      |||||||
XX      17 aArgLYsGlnLeuValProLeuLeuTYrLYsGlnPhenValProGlyValP 34
XX      |||||||
XX      108 GCGCAGACAGCTCTGCGCGCTACTCTACACACATTGTGCGCGCGGCG 157
XX      |||||||
XX      34 roGluArGTrLeuGlyAlaSerGlyProAlaGluGlyArGValAlaArg 50
XX      |||||||
XX      158 CAGAGCGACCTCTGGCGCGCGCGCGCGCGAGGAGGAGCGTGGCAAG 207
```

```
51 GlySerGluArGpheaArgAspLeuValProAsnTYrAsnProAspIleI 67
|||
208 GGCTCCGAGCGCTCCGGACCTCGTGCCCAACTACAACCCCGACATCAT 257
67 ePhelYsAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||
258 CTTCAAGATGAGGAGAACACTGGAGCCGCGCTATGACCGACGCTT 307
84 ySLySGluArGValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||
308 GCAAGAGAGAGGTGAACGCTTGCCCACTGCCGTGATGACATATGCGCC 357
101 GLyValArGLeuArGValThrGluGlyTYrAspGluAspGlyHisHisAl 117
|||
358 GGAGTGGCGCTACGAGTGAAGGCTGGGAGGAGGAGGAGGAGGAGGAGC 407
117 acLInAspSerLeuHisTYrGluGlyArgAlaLeuAspIleThrIsera 134
|||
408 TCAGGATTCACCTCCACTACGAGGCCGCTTGTTGGACATCACTACGCTG 457
134 sPArGAsPArGAsnLYsTYrGlyLeuLeuAlaArGLeuAlaValAla 150
|||
458 ACCGCGACCGGACACAGATGCGTGGCGCGCGCTCGCAGTGAAGCC 507
151 GLyPheaSPTrpValTYrTYrGluSerArGAsnHisValHisValSera 167
|||
508 GGCTTCGACTGGGTCTACTACGAGTCCCGCACCACGTCAGTGGCT 557
167 lLYsAlaAspAsnSerLeuAlaValArGAlaGlyGly 179
|||
558 CAAGCTGATACCTCACTGCGGCGCGCGCGCGCGCG 594
seq_name: /STDS2/gc9data/geneseq/geneseqn/NA2000.DAT:252264
seq_documentation_block:
ID 252264 standard; DNA; 1190 BP.
XX
XX 252264:
XX
XX 18-JUL-2000 (first entry)
XX
XX Human Desert hedgehog (Dhh) gene.
XX
XX Desert hedgehog gene; DhH; human; growth modulator; therapeutic agent;
XX lung; hedgehog; hh; patched; ptc; fibroblast growth factor; fgf-10;
XX antiproliferative; anticancer; vulnerrary; antineumatic; hypotensive;
XX anti-inflammatory; antiaesthetic; antiarthritic; tuberculostatic; asthma;
XX antimicrobial; antiallergy; treatment; prevention; lung diseases; cancer;
XX cystic fibrosis; bronchitis; emphysema; respiratory distress syndrome;
XX tuberculosis; wound healing; lung transplantation; ds.
XX
XX Homo sapiens.
XX
XX OS
XX FH
XX Key 1..1189 Location/Qualifiers
XX CDS
XX
XX /*tag= a
XX /*product= "Human Desert hedgehog (Dhh) protein"
XX /*note= "Does not include stop codon"
XX /partial
XX
XX WO200015246-A2.
XX
XX 23-MAR-2000.
XX
XX 10-SEP-1999: 99WO-US20500.
XX
XX 11-SEP-1998: 98US-0099952.
XX
XX (HARD ) HARVARD COLLEGE.
XX
XX Peptocell C, Lewis P, McMahon AP:
```

DR WPI: 2000-271252/23.
P-PSDB: Y70683.

XX Modulation of lung tissue or cell growth rate useful for treating or
PT preventing damage to lung tissue comprises ectopically contacting
PT tissue with hedgehog therapeutic, patched therapeutic or fibroblast
PT growth factor-10 -

XX Claim 15: Page 106-109; 143pp; English.

XX The patent discloses a method for modulating the growth state of
CC epithelial or mesenchymal cells of the lung, by ectopically contacting
CC the tissue with a therapeutic agent, that can effectively alter the rate
CC of proliferation of cells. This agent can be selected from hedgehog (hh),
CC patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It
CC involves a direct or indirect antagonism of patched-mediated regulation
CC of gene expression. This method is useful for the treatment or prevention
CC of lung diseases, like cancer, cystic fibrosis, bronchopneumocystosis,
CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,
CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress
CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary
CC pulmonary hypertension. It is also used to control wound healing or other
CC reformation processes in the lung and augment lung transplantation. The
CC present DNA sequence is the human desert hedgehog (Dhh) gene, essential
CC for development of the respiratory system. Hedgehog polypeptides can be
CC used to control the formation and/or maintenance of the lung tissue.

XX Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment_scores:
Quality: 950.00 Length: 179
Ratio: 5.307 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-08-900-220C-17_COPY_20_198 x 252264 ..

Align seg 1/1 to: 252264 from: 1 to: 1190

1 AAGInserCysGlyProGlyArgGlyProValGlyArgArgGlyTyrAl 17
|||||
58 GCGCAGAGCTGCGGGCGGGCGGGCGGCGGTGGCGGGCGCGCATATC 107
17 aargtysgInLeuValProLeuLeuTyrTysInPheValProGlyValP 34
|||||
108 GCGCAGAGCAGCTGCGCCGCTACTCTACAGCAATTTGCGCGCGCTGC 157
34 rogluarGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CAGACCGGACCTGGCGCCGACGTCGAGCGAGCGGCGGCGGCGCAAGC 207
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspLeu 67
|||||
208 GCGTCGAGCGCTTCGCGGACCTCGGCCCACTCAACCCGACGACTCAT 257
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAAGGATGAGGAACAGTGAAGCCGACCGCTGATGACGAGCGCTT 307
84 yLysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTyrPro 100
|||||
308 GCAGAGAGAGCGGTGAACGCTTGGCCATGCGGTATGAACATGTGCGCC 357
101 GtValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAl 117
|||||
358 GGAAGTCCGCTTACGAGTGAAGGCTGGAGCGAGAGCGGCGCACCAAGC 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTyrSera 134
|||||
408 TCAGGATTCATCTCACTACGAGAGCGGCTGTTGACATCACTACAGCTTC 457
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150

|||||
458 ACCGCGACCGCAACAGTATGGTTGCTGCGCGGCTCGACGTGAAGCC 507
151 GlyPheAspTyrValTyrGluSerArgAsnHisValHisValSera 167
|||||
508 GCGTCGAGCTGGGTCTACTACAGAGTCCGCAACGACGTCACGTGCTGGT 557
167 lLysAlaAspAsnSerLeuAlaValAlaArgAlaGlyGly 179
|||||
558 CAAGCTGATTAACCTCACTGCGGTCGCGGCGGCGCGC 594

seq_name: /STS2/gcdata/geneseq/geneseqn/NA1999.DAT:200317

seq_documentation_block:

ID 200317 standard; DNA: 1820 BP.

XX 200317;

DT 22-OCT-1999 (first entry)

XX Human desert hedgehog (Dhh-H) protein encoding DNA.

DE Desert hedgehog; Dhh-H; male infertility; menopause; human; CNS;

KW spinal growth injury; spinal cord regeneration; Parkinson's disease;

KW tumour growth; central nervous system disorder; Alzheimer's disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 180..1370

FT /tag= a /transl_except= (pos:438..440, aa:Xaa)

FT /product= "Dhh-H protein"

FT /note= "Xaa = unknown"

FT sig_peptide 180..245

FT /tag= b

FT /note= "putative signal sequence"

FT mat_peptide 246..1367

FT /tag= c

XX W09939725-A1.

XX 12-AUG-1999.

XX 04-FEB-1999; 99WO-US02440.

XX 06-JUL-1998; 98US-0091843.

XX 06-FEB-1998; 98US-0073878.

XX 02-MAR-1998; 98US-0076553.

XX (ELIL) LILLY & CO ELI.

XX Edmonds BT, King MW, Smith RC;

XX WPI: 1999-508454/42.

XX P-PSDB: Y27264.

XX New human desert hedgehog (Dhh-H) protein useful for, e.g. treating

XX male infertility and menopause

XX Example 1: Page 58-61; 69pp; English.

XX This DNA encodes a human desert hedgehog (Dhh-H) protein. The Dhh-H
CC protein can be expressed by standard recombinant methodology.

CC Compositions containing dhh-H form methods of treating male infertility
CC and menopause in humans. The compositions are additionally useful for

CC treating spinal growth injury, promoting spinal cord regeneration,
CC inhibiting tumour growth and treating central nervous system (CNS)

CC disorders in humans. Such CNS disorders includes Alzheimer's disease
CC and Parkinson's disease. Dhh-H is also useful for identifying compounds that

CC bind to the protein. The use of Dhh-H for producing antibodies that are

XX useful in diagnosis and therapy.

SQ Sequence 1820 BP; 336 A; 520 C; 620 G; 343 T; 1 other;

alignment_scores:
 Quality: 950.00 Length: 179
 Ratio: 5.307 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-900-220C-17_COPY_20_198 x 200317

Align seg 1/1 to: 200317 from: 1 to: 1820

```

1  Alaglnsercysgilyproglyargglyprovalglyargargtyr 17
  |||||||
237 GCCGAGAGCTGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGG 286
  |||||||
17  aarglysglnleuvalproleuleutyrylsglnphevalprogly 34
  |||||||
287 GCCGACAGCTGCTGGCGCTACTCTACAGCAATTGTGCCCCGGCTGC 336
  |||||||
34  rogluargthrleuglyalaserglyproalaglulglyargvala 50
  |||||||
337 CAGAGCGGACCTGGCGCCAGTGCGCAGCGAGGGAGGGGTGCCAAG 386
  |||||||
51  glysergluargpheargaspleuvalproasntyrsnproasp 67
  |||||||
387 GCGTCCAGCGCTTCGGGACCTGTCGCCCAACACACCCGACATCAT 436
  |||||||
67  ephelysaspcluglunserserglyalaaspargleumethr 84
  |||||||
437 CTTYAAGAGTAGAGAAACAGTGGAGCCGCGCTGATGCCAGCGTT 486
  |||||||
84  ysllysgluargvalasnalaleualallealavalmetasmet 100
  |||||||
487 GTAAGGAGCGGGTGAACGCTTGGCCATTTGCCGTGATGAACATG 536
  |||||||
101 glyvalargleuargvalthrlyglutraspplunaspglyhsh 117
  |||||||
537 GGAGTGGCGCTACGAGTACTGAGCGCTGGGACGAGGAGCGCCAC 586
  |||||||
117 aglnaspserserleuistyrlyglulglyargalaaleuaspl 134
  |||||||
587 TCAGATTTCACTCCACTACGAGAGCGCGTGTGACATCACTAC 636
  |||||||
134 spargaspargasnlystyrelyleuleualargleualaval 150
  |||||||
637 ACCCGACCGCAACAGTATGGGTGCTGGCGCGCTCGCAGTGA 686
  |||||||
151 glypheasprvaltyrtyrlygluserfargasnhsvalhsval 167
  |||||||
687 GGCCTTCAGCTGGGTCTACTACGAGTCCCGCAACACGCTCCAG 736
  |||||||
167 llysalaspsasnserleualavalargalagly 179
  |||||||
737 CAAGCTGATTAACACTGCGGCTCCGGCGCGGGCGGC 773
  |||||||

```

seq_name: /SID32/gcdata/geneseq/geneseqn/NA2000.DAT.A58931

seq_documentation_block:

ID A58931 standard; DNA; 1820 BP.

XX A58931;

XX AC 20-OCT-2000 (first entry)

XX DE DNA encoding a desert hedgehog protein.

XX XX Human; desert hedgehog protein; Dhh-H; male menopause; male virility;

KW testostosterone level; male infertility; erectile dysfunction;

KW tumour growth; spinal regeneration; ss.

XX XX Homo sapiens.

XX Key Location/Qualifiers
 FH CDS 180..1370
 FT /*tag= a
 FT /product= "desert hedgehog protein"
 PN WO200041537-A2.

XX 20-JUL-2000.

XX 12-JAN-2000; 2000MO-US00706.

XX 14-JAN-1999; 99US-0115931.

XX (ELIL) LILLY & CO ELI.

XX Edmonds BT;

XX MPI; 2000-475908/41.

XX P-PDB; B07528.

XX PT Method of treating male menopause, increasing testosterone levels and

XX PT enhancing male virility uses administration of a human homolog of

XX PT desert hedgehog protein -

XX PS Example 1; Page 45-49; 52pp; English.

XX CC The present sequence encodes a human desert hedgehog protein (Dhh-H).

XX CC The polypeptide is used for treating male menopause. Dhh-H is used

XX CC for increasing testosterone levels and enhancing male virility.

XX CC Dhh-H can also be used to produce antibodies which are then useful for

XX CC diagnosis to screen for Dhh-H or potential modulators of Dhh-H e.g. in

XX CC competitive displacement assay or in therapeutic applications.

XX CC Dhh-H protein is used to treat male infertility, male menopause,

XX CC erectile dysfunction, diminished virility, tumour growth or spinal

XX CC regeneration.

SQ Sequence 1820 BP; 336 A; 520 C; 620 G; 343 T; 1 other;

alignment_scores:
 Quality: 950.00 Length: 179
 Ratio: 5.307 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-08-900-220C-17_COPY_20_198 x A58931

Align seg 1/1 to: A58931 from: 1 to: 1820

```

1  Alaglnsercysgilyproglyargglyprovalglyargargtyr 17
  |||||||
237 GCCGAGAGCTGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGG 286
  |||||||
17  aarglysglnleuvalproleuleutyrylsglnphevalprogly 34
  |||||||
287 GCCGACAGCTGCTGGCGCTACTCTACAGCAATTGTGCCCCGGCTGC 336
  |||||||
34  rogluargthrleuglyalaserglyproalaglulglyargvala 50
  |||||||
337 CAGAGCGGACCTGGCGCCAGTGGCGCAGCGGAGGGAGGGGTGCCA 386
  |||||||
387 GCGTCCAGCGCTTCGGGACCTGTCGCCCAACACACCCGACATCAT 436
  |||||||
67  ephelysaspcluglunserserglyalaaspargleumethr 84
  |||||||
437 CTTYAAGAGTAGAGAAACAGTGGAGCCGCGCTGATGCCAGCGTT 486
  |||||||
84  ysllysgluargvalasnalaleualallealavalmetasmet 100
  |||||||
487 GTAAGGAGCGGGTGAACGCTTGGCCATTTGCCGTGATGAACATG 536
  |||||||

```


XX Human Desert hedgehog protein N-terminal region cDNA.
 XX Desert hedgehog; HUDHH; human; ds.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT sig_peptide 7..72 /*tag- a
 FT mat_peptide 73..600 /*tag- b
 FT
 XX EP874048-A2.
 XX
 XX PD 28-OCT-1998.
 XX
 XX PF 24-APR-1998: 98BP-0303187.
 XX
 XX PR 14-APR-1998: 98BP-0117873.
 XX 25-APR-1997: 97JP-0121578.
 XX
 XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 XX PI Ariyasu T, Nakamura S, Orita K;
 XX WPI; 1998-544642/47.
 XX DR P-PSDB; W79597.
 XX
 XX Human Desert hedgehog protein - and corresponding DNA and monoclonal
 PT antibody
 XX
 PS Example 1-2: Page 27-28; 39pp; English.
 XX
 CC This cDNA sequence codes for the N-terminal region (see W79597) of
 CC a precursor form (see W79595) of novel human Desert hedgehog protein
 CC (see also W79593). Human Desert hedgehog protein DNA (HUDHH) was
 CC originally identified in human following ARH-77 (ATCC CRL-1621) cell
 CC line following RT-PCR amplification screenings. Further PCR (see
 CC V62405-06) yielded cDNA encoding the N-terminal region. DNA
 CC molecules (see V62393-95) encoding mature and precursor forms (see
 CC W79593-95) of human Desert hedgehog are claimed, as are a monoclonal
 CC antibody (Mab) that recognises the protein, a process for producing
 CC the protein, and a method for detecting the protein using the Mab.
 CC The hedgehog protein, DNA and Mab can be used to elucidate
 CC hereditary morphological abnormalities in humans to establish their
 CC treatments and diagnoses.
 CC
 SO Sequence 602 BP; 111 A; 187 C; 199 G; 105 T; 0 other;
 alignment_scores: Quality: 949.00 Length: 179
 Ratio: 5.302 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.441
 alignment_block:
 US-08-900-220C-17_COPY_20_198 x V62397 ..
 Align seg 1/1 to: V62397 from: 1 to: 602
 1 AAGTnSerGlyProGlyArgGlyProValGlyArgArgGlyTyrAl 17
 64 GCCCAGAGCTCGGGCCGGGGCGGGTGGCCGCCGCCCTATGC 113
 17 AAGTnSerGlyProGlyArgGlyProValGlyArgArgGlyTyrAl 34
 114 GCCCAGAGCTCGGGCCGGGGCGGGTGGCCGCCGCCCTATGC 163
 34 TGGTnSerGlyProGlyArgGlyProValGlyArgArgGlyTyrAl 50
 164 CAGAGCGGAGCTCGGGCCGGGGCGGGTGGCCGCCGCCCTATGC 213

51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspGleI 67
 214 GGCCTCCGAGCGCTTCCGGGACCTCGTCCCAACTCAACCCGACATCAT 263
 67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
 264 CTTCAGGATGAGGAGAACAGTGGAGCCGCGCTGATGACCGAACCTT 313
 84 YSLysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
 314 GTAAGGAACGGGTGAACGCTTGGCCATTTGCCGTGATGACATGAGGCC 363
 101 GlyValArgLeuArgValThrGluGlyTrpAspGluAspGlyHisHisAl 117
 364 GAGAGCGCCCTACGAGTACGAGGCGTGGAGCAGGACCGCCACACGCG 413
 117 AGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspLeuThrTrpSerA 134
 414 TCAGGATTCACCTCCACTACGAAAGCGGCTTTGGACATCCTACGCTG 463
 134 SPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
 464 ACCGCGACCGCAACAAGTATGCGTTGCTGGCGCGCTCGCAGTGAAGCC 513
 151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVa 167
 514 GGCCTCGAGCTGGGTCTACTACGAGTCCCGCAACCAATCCAGTGTCCGT 563
 167 LlysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
 564 CAAAGCTGATNACTCTGCGGCTCGGGCGGCGGCGC 600
 seq_name: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT.V62395
 seq_documentation_block:
 ID V62395 standard; cDNA; 1188 BP.
 XX
 AC V62395:
 XX
 DT 02-FEB-1999 (first entry)
 XX
 DE Human Desert hedgehog protein cDNA.
 KW Desert hedgehog; HUDHH; human; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT sig_peptide 1..66 /*tag- a
 FT mat_peptide 67..594 /*tag- b
 FT
 XX EP874048-A2.
 XX
 XX PD 28-OCT-1998.
 XX
 XX PF 24-APR-1998: 98BP-0303187.
 XX
 XX PR 14-APR-1998: 98BP-0117873.
 XX 25-APR-1997: 97JP-0121578.
 XX
 XX PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 XX
 XX PI Ariyasu T, Nakamura S, Orita K;
 XX WPI; 1998-544642/47.
 XX DR P-PSDB; W79595.
 XX
 XX Human Desert hedgehog protein - and corresponding DNA and monoclonal
 PT antibody
 XX
 PS Claim 10: Page 24-25; 39pp; English.

XX This cDNA sequence codes for a precursor (see W79593) of a novel
 CC human Desert hedgehog protein (see also W79593). Human Desert
 CC hedgehog protein DNA (HuHh) was originally identified following
 CC screenings of human cell lines by PCR using primers (see V62402-03)
 CC based on the mouse Desert hedgehog gene. These screenings
 CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is
 CC derived from the plasma cell of a leukemia patient, expressed a
 CC specific gene at an elevated level. Sequencing confirmed it to
 CC be a novel human gene showing homology to the mouse Desert hedgehog
 CC gene. DNA encoding N-terminal sequences (see V62397) was obtained
 CC by further PCR amplifications (see V62404-06). The invention
 CC provides Desert hedgehog polynucleotides and protein, a monoclonal
 CC antibody (Mab) that recognises the protein, a process for producing
 CC the protein, and a method for detecting the protein using the Mab.
 CC The hedgehog protein, DNA and Mab can be used to elucidate
 CC hereditary morphological abnormalities in humans to establish their
 CC treatments and diagnoses.
 XX
 S0 Sequence 1188 BP; 179 A; 376 C; 419 G; 214 T; 0 other;

alignment_scores:
 Quality: 949.00 Length: 179
 Ratio: 5.302 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.441

alignment_block:
 US-08-900-220c-17_COPY_20_198 x V62395 ..

Align seg 1/1 to: V62395 from: 1 to: 1188

```

1  AAlaInserCySgLyProGlyArgGlyProValGlyArgArgGlyTyrAl 17
  |||||||
58  GCCCAGAGCTGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 107
  |||||||
17  AArgLySgInLeuValProLeuLeuTyrLySgInPheValProGlyValP 34
  |||||||
108  GCGCAAGAGCTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157
  |||||||
34  roGluArgThrLeuGlyAlaSerGlyProAlaGlyGlyArgValAlaArg 50
  |||||||
158  CAGACCGGAGCCTGGGCGCCAGTGGGCGGAGGAGGAGGAGGAGGAGG 207
  |||||||
51  GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspLeu 67
  |||||||
208  GGCTCGGAGCGCTTCGGGAGCTCGGCGCCACTGACCAACCCGACATCAT 257
  |||||||
67  ePheLySAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
  |||||||
258  CTTCAAGATGAGGAGAACAGTGGAGCGCGCTGATGACCGCAACGCTT 307
  |||||||
84  yLySgLyArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
  |||||||
308  GTAAAGAACGGGTGAACGCTTTGGCCATTGCGGTGATGAACATGTGGCCC 357
  |||||||
101  GLyValAlaArgLeuArgValThrGlyTyrPaspGluAspGlyHisHisAl 117
  |||||||
358  GGAGTGGCGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 407
  |||||||
117  aGlnAspSerLeuHisTyrGlyGlyArgAlaLeuAspIleThrTrpSerA 134
  |||||||
408  TCAGGATTCACCTACACAGAGCGCGCTTTGGACATCCTACCTGCTG 457
  |||||||
134  sPArAspArgAsnLySgLyLeuLeuAlaArgLeuAlaValGluAla 150
  |||||||
458  ACCGGAGCGGCAACAAGTATGGTGTGGTGGTGGTGGTGGTGGTGGTGG 507
  |||||||
151  GLyPheAspTrpValTyrTyrGlyGluSerArgAsnHisValHisValSerVa 167
  |||||||
508  GGCTTCGACTGGCTGCTACTAGAGTCCCGCAACACATCCTGCTGCTGCT 557
  |||||||
167  lLySAlaAspAsnSerLeuAlaValAlaArgAlaGly 179
  |||||||

```

seq_name: /SIB2/gcdata/geneseq/geneseq/NA1998.DAT:V62393

seq_documentation_block:
 ID V62393 standard; cDNA; 528 BP.

AC V62393;
 DT 02-FEB-1999 (first entry)

DE Human Desert hedgehog mature protein cDNA.

KW Desert hedgehog; HuHh; human; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH mat_peptide 1..528
 FT /*tag= a

PN EP874048-A2.

PD 28-OCT-1998.

PF 24-APR-1998; 98EP-0303187.

PR 14-APR-1998; 98JP-0117873.

PR 25-APR-1997; 97JP-0121578.

PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

PI Ariyasu T, Nakamura S, Orita K;

DR WPI: 1998-544642/47.

DR P-PSDB: W79593.

PT Human Desert hedgehog protein - and corresponding DNA and monoclonal antibody

PS Claim 8; Page 20-21; 39pp; English.

XX This cDNA sequence codes for novel human Desert hedgehog protein
 CC mature polypeptide (see W79593). Human Desert hedgehog protein
 CC DNA (HuHh) was originally identified following extensive
 CC screenings of human cell lines by PCR using primers (see V62402-03)
 CC based on the mouse Desert hedgehog gene. These screenings
 CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is
 CC derived from the plasma cell of a leukemia patient, expressed a
 CC specific gene at an elevated level. Sequencing confirmed it to
 CC be a novel human gene showing homology to the mouse Desert hedgehog
 CC gene. Nucleotide sequences (see V62394-95) encoding precursor
 CC forms (see W79594-95) of human Desert hedgehog are also claimed,
 CC as are a monoclonal antibody (Mab) that recognises the protein, a
 CC process for producing the protein, and a method for detecting the
 CC protein using the Mab. The hedgehog protein, DNA and Mab can be
 CC used to elucidate hereditary morphological abnormalities in humans
 CC to establish their treatments and diagnoses.

S0 Sequence 528 BP; 101 A; 160 C; 181 G; 86 T; 0 other;

alignment_scores:
 Quality: 936.00 Length: 176
 Ratio: 5.318 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.432

alignment_block:
 US-08-900-220c-17_COPY_20_198 x V62393 ..

Align seg 1/1 to: V62393 from: 1 to: 528


```

137 rgasnlystygylleuleualaargleualaValaGluAlaGlypheasp 153
|||||
401 GCAACAGATATGGCTTGGCGCCCTCGAGTGAAGCCGGCTTCGAC 450
154 TTPVALTYTYRGLuSerArgasnHisValHisValSerVallyAlaas 170
|||||
451 TGGCTACTACGAGTCCCGCAACCAATCCACGTGCGTCAAGCTCA 500
170 PAsnSerleualaValaArgAlaGlyly 179
|||||
501 TAACCTACTGGCGGTCCGGCGCGCGGC 528

seq_name: /SIDS2/gcdata/geneseq/geneseqn/NA1995.DAT:Q91642
seq_documentation_block:
ID Q91642 standard; cDNA; 1190 BP.
XX
XX Q91642;
XX
XX 18-MAR-1996 (first entry)
XX
DE Mouse desert hedgehog protein gene.
XX
KM Mouse; desert hedgehog gene; probe; primer; diagnostic;
KM nervous system disorder; gene therapy; antibody; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT exon 1..303
FT /*tag= a
FT exon 304..567
FT /*tag= b
FT exon 568..1188
FT /*tag= c

WO9518856-A1.
13-JUL-1995.
XX
PF 30-DEC-1994; 94MO-US14992.
XX
PR 14-DEC-1994; 94US-0356060.
PR 30-DEC-1993; 93US-0176427.
XX
PA (HARD ) HARVARD COLLEGE.
PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.
XX
PI Tingham PW, Mcmahon AP, Tabin CJ;
XX
DR WPI: 1995-255060/33.
DR P-PSDB: R77345.
XX
PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful
PT to treat degenerative nervous system disorder(s) and in gene
PT therapy.
XX
PS Claim 5; Page 135-37; 210pp; English.
XX
CC The sequence encodes a mouse desert hedgehog protein, homologous
CC to a Drosophila hedgehog protein (R77337), and has been isolated by
CC low stringency screening of a mouse genome DNA library, in phage
CC lambda. The sequence contains 3 homologous regions, encoding a
CC single open reading frame interrupted by introns. Splicing has
CC been confirmed by polymerase chain reaction amplification of first
CC strand cDNA generated from adult testicle RNA. Probes and primers
CC derived from hedgehog sequences may be used as diagnostic agents
CC for neuromuscular, autonomic or central nervous system disorders,
CC and the gene may also be used in gene therapy. Antibodies
CC generated from the encoded protein may be used as therapeutic or
CC research reagents.
XX

```

```

50 Sequence 1190 BP; 194 A; 371 C; 399 G; 226 T; 0 other;
alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:
US-08-900-220c-17_copy_20_198 x Q91642 ..
Align seg 1/1 to: Q91642 from: 1 to: 1190

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgTyrAl 17
|||||
58 GCCACAGCTGCGGGCGCGGCGGAGACCGGTTGGCGGGCGGCTATATG 107
17 ArgGlyGlnLeuValProLeuLeuTyrTyrGlnPheValProGlyValP 34
|||||
108 GCGCAAGCAACTTGTGCTCTGTATACAAGCAGTTTGTGCCACGATGCG 157
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CCGAGCGGACCTCGGCGCGGAGTGGCCGACGAGGAGGAGGATTAACAAG 207
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
208 GGGTCGAGCGGCTTCGGGACCTCGTACCCAACTACACCCGACATATAT 257
67 ePheLysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
258 CTTCAGAGATGAGAGAGACAGCGCGCCAGACCCCTGATGACAGAGCGGT 307
84 YstLysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTPro 100
|||||
308 GCAAGACCGCGGTGAACCTCTAGCCATCCGGGTGATGAACATGTGGCCC 357
101 GlyValArgLeuArgValThrGluGlyTTPAspGluAspIleHisIAsI 117
|||||
358 GGAGTAGCGCTACGTGTGACTGAAAGCTGGAGGAGGAGGCCACCAACGC 407
117 ArgAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThSerA 134
|||||
408 ACAGGATTCACTCCACTACGAAGCCGCTGTGGACATCACACACGCTGTG 457
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGTGACCGTAAATAGTATGTTTGTGGCGCGCTAGCTGTGGAAGCC 507
151 GlyPheAspTyrValTyrTyrGluSerArgAsnHisValHisValSerVa 167
|||||
508 GGATTGACGTGGGTCTACTAGAGTCCCGCAACCAATCCACGTATACGCT 557
167 llyAlaAspAsnSerLeuAlaValaArgAlaGlyly 179
|||||
558 CAAAGCTGATTAACCTACGTGCGGTCCGAGCGGAGGC 594

```

THIS PAGE BLANK (USPTO)

OM of: US-08-900-220c-17_COPY_20_198 to: GenEmbl:* out_format : pfs
Date: Apr 22, 2001 12:04 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame+g2n.model -DEV=xlp  
-O/cgn2_1/USPTO_spool/US08900220/runat_19042001_104047_21668/app-query.fasta_1.253  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -ICOPT=0.000 -ICOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -EGAPOP=6.000  
-EGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=2000000000 -USRR=US08900220_@CGNL_1_4706 -NCPU=6  
-ICPU=3 -LONLOG -NO_XLPPX -WAIT -THREADS=1
```

Search information block:

```
Query: US-08-900-220c-17_COPY_20_198  
Query length: 179  
Database: GenEmbl:*  
Database sequences: 1283235  
Database length: -1216004940  
Search time (sec): 1235.880000
```

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_pat1:AX054868	+	950.00	1409.25	1190	AX054868 Sequence 15 from Pat
gb_pat1:AR021199	+	932.00	1382.48	1190	AR021199 Sequence 3 from Paten
gb_pat1:AX053081	+	932.00	1382.48	1190	AR063081 Sequence 2 from Paten
gb_pat1:AX054856	+	932.00	1382.48	1190	AX054856 Sequence 3 from Paten
gb_r01:MM0HH	+	932.00	1382.48	1191	X76292 M.musculus (129/Su) Dht
gb_ov:XL026349	+	823.50	1221.12	1191	U26349 Xenopus laevis cephalic
gb_r01:MM085610	+	809.00	1199.51	1197	U26350 Xenopus laevis hedgehog
gb_pat1:AX054866	+	743.50	1097.99	6.1e-53	U85610 Mus musculus Indian hed
gb_pat1:AR021198	+	741.00	1097.91	6.2e-53	AR021198 Sequence 1 from Paten
gb_pat1:AR063080	+	741.00	1097.91	6.2e-53	AR063080 Sequence 1 from Paten
gb_pat1:AX054854	+	741.00	1097.91	6.2e-53	AX054854 Sequence 1 from Paten
gb_ov:CHKLPAMD	+	741.00	1096.42	7.5e-53	L28099 Gallus gallus sonic hed
gb_pat1:AX054866	+	740.50	1095.42	8.5e-53	AX054866 Sequence 13 from Pat
gb_pat1:AR021201	+	740.00	1096.22	7.3e-53	AR021201 Sequence 7 from Paten
gb_pat1:AX054860	+	740.00	1096.22	7.6e-53	AX054860 Sequence 4 from Paten
gb_pat1:AX063083	+	738.00	1092.65	1.2e-52	AX063083 Sequence 7 from Paten
gb_r01:MM0HH	+	738.00	1091.91	1.3e-52	L38518 Homo sapiens sonic hed
gb_pat1:AX054864	+	737.00	1091.75	1.4e-52	AX054864 Sequence 11 from Pat
gb_r01:MM0HH	+	734.50	1085.12	3.2e-52	U85511 Gallus gallus Indian he
gb_ov:GGU58511	+	723.00	1081.99	4.7e-52	D63339 Cynops pyrrhogaster son
gb_pat1:AX054858	+	723.00	1080.79	5.5e-52	AX054858 Sequence 5 from Paten
gb_pat1:EL16679	+	723.00	1086.48	2.7e-52	EL16679 DNA encoding hedgehog p
gb_pat1:AR021202	+	723.00	1071.26	1.9e-51	AR021202 Sequence 5 from Paten
gb_pat1:AR063084	+	723.00	1071.26	1.9e-51	AR063084 Sequence 5 from Paten
gb_pat1:AX054862	+	723.00	1071.26	1.9e-51	AX054862 Sequence 9 from Paten
gb_ov:DR030711	+	723.00	1069.12	2.5e-51	U30711 Danio rerio sonic hedge
gb_ov:BRSONHP	+	723.00	1065.93	3.7e-51	L23669 B.rerio shh mRNA for sc
gb_ov:ZEFVH1A	+	721.00	1068.32	2.7e-51	L27585 Brachydanio rerio (vhl-1
gb_ov:DR030710	+	721.00	1068.32	2.7e-51	U30710 Danio rerio tigris-wiki
gb_pat1:AX054870	+	718.00	1061.55	6.5e-51	L27340 Rat (vhl-1) mRNA, comp
gb_r01:RATVNH1A	+	716.00	1060.41	7.6e-51	U26314 Xenopus laevis sonic he
gb_ov:XL026314	+	716.00	1059.34	8.7e-51	L35248 Xenopus laevis sonic he
gb_ov:AB029748	+	716.00	1058.37	9.6e-51	AB029748 Paralichthys olivaceu
gb_ov:ELKSH	+	715.00	1055.98	1.3e-50	L35213 Xenopus laevis morphog
gb_ov:AF047466	+	713.00	1056.54	1.2e-50	AF047466 Notophthalmus virides
gb_ov:XL026404	+	713.00	1053.37	1.9e-50	U26404 Xenopus laevis banded h
gb_ov:AF003532	+	707.00	1041.20	8.9e-50	AF003532 Pleurodeles waltl son
gb_int:AB044709	+				AB044709 Gyrillus blinaculatus m

```
gb_ov:DRECHINDA + 704.00 1041.81 8.2e-50 1479 1 Y08426 D.rerio mRNA for EHR  
gb_r01:AF148226 + 660.00 986.18 1.0e-46 386 1 AF148226 Rattus norvegicus d  
gb_int:BFAMPHIHH + 640.00 940.12 3.8e-44 3610 1 Y13858 Branchiostoma florida  
gb_ov:AF144100 + 618.00 923.15 3.3e-43 417 1 AF144100 Bos taurus sonic he
```

seq_name: gb_pat1:AX054868

seq_documentation_block:

```
LOCUS AX054868 1190 bp DNA PAT 13-JAN-2001  
DEFINITION Sequence 15 from Patent WO0073337.  
ACCESSION AX054868  
VERSION AX054868.1 GI:12228297  
KEYWORDS  
SOURCE human.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
```

REFERENCE

1 (bases 1 to 1190)

AUTHORS

Peplinsky,R.B., Taylor,F. and Garber,E.

TITLE

Polymer conjugates of hedgehog proteins and uses

JOURNAL

Patent: WO 0073337-A 15 07-DEC-2000;
BIOGEN, INC. (US)

FEATURES

Location/Qualifiers

source

1..1190
/organism="Homo sapiens"
/db_xref="taxon:9606"

```
BASE COUNT 176 a 375 c 424 g 215 t  
ORIGIN
```

alignment_scores:

```
Quality: 950.00 Length: 179  
Ratio: 5.307 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000
```

alignment_block:

US-08-900-220c-17_COPY_20_198 x AX054868 ..

Align seg 1/1 to: AX054868 from: 1 to: 1190

```
1 AAGlnSerCygGlyProGlyArgGlyProValAGlyArgArgTyrAl 17  
|||||  
58 GCCCAGAGCTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGCTATGC 107  
|||||  
17 AArglySGlnLeuValProLeuLeuTyrTrpGlnPheValProGlyValP 34  
|||||  
108 GCGCAGCGAGCTCGCGCGCTCTCTACAGCAATTTGGCCGGCGGCTGC 157  
|||||  
34 roGluArgThrLeuGlyAlaSerGlyProAlaGlyGlyArgValAlaArg 50  
|||||  
158 CAGAGCGGAGCCTGGCGCGCGAGTGGCGCGAGCGGAGCGGCGGCGCAGG 207  
|||||  
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67  
|||||  
208 GGCCTCGAGCGCTCCGCGAGCTCGTGGCCCACTACACCCCGCATCAT 257  
|||||  
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84  
|||||  
258 CTTCAAGAGTGAAGAGACAGTGGACCCAGCCGCTGATGACCGAGCTT 307  
|||||  
308 GCAAGAGAGAGGGTGAACGTTTGGCCATTCGCGGTGTAACATGTCGCC 357  
|||||  
101 GtValArgLeuArgValThrGluGlyTyrPArgGluAspGlyHisAla 117  
|||||  
358 GAGAGTGGCTAGTGAAGTGAAGGCTGGAGCGAGGAGCGGACCGACCGC 407  
|||||  
117 AGlnAspSerLeuHisTyrGluGlyAlaGlyAlaLeuAspIleThrSerA 134  
|||||  
408 TCAGAGTATACGTCACCTAGAGAGCGGCTGTTGGACATCATAGTCTGCG 457  
|||||  
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
```

|||||
458 ACCGCGACCCCAACAGTATGGGTGCTGGCGCGCTCGACATGGAAGCC 507
151 GLYPheSPTRPValTYrTYrGLuSerARgAsnHISValHISValSerVa 167
|||||
508 GGCTTCACAGTGGGTCTACTACGAGTCCCGCACACGACGTCCACGTGGGT 557
167 lYsAlAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||
558 CAAAGCTGATTAACACTGCGCGGTCCGCGCGGCGCGC 594

seq_name: gb_pat1:AR021199

seq_documentation_block:
LOCUS AR021199 1190 bp DNA PAT 05-DEC-1998
DEFINITION Sequence 3 from patent US 5789543.
ACCESSION AR021199
VERSION AR021199.1 GI:3975814
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1190)
AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins and uses related thereto
JOURNAL Patent: US 5789543-A 3 04-AUG-1998;
FEATURES
source 1. 1190 Location/Qualifiers
BASE COUNT 194 a 371 c 399 g 226 t
ORIGIN

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:

US-08-900-220C-17_COPY_20_198 x AR021199 ..

Align seg 1/1 to: AR021199 from: 1 to: 1190

1 AlaGInSerCySGlyPProGlyARgGlyProValGlyARgARgTYrAl 17
|||||
58 GCCGAGAGCTGCGGGCGCGGCGGAGACCGGTGGCGCGCGCTTATGT 107
17 aArgLYGlnLeuValProLeuLeuTYrLYSGlnPheValProGlyValP 34
|||||
108 GCGCAACCACTTGCTGCTGCTATACAAAGCAAGTTGTGCCAGTATGC 157
34 roGluARgThrLeuGlyAlaSerGlyProAlaGluGlyARgValAlaArg 50
|||||
158 CCGAGCGGACCTTGCGCGAGTGGCCAGCGAGGGAGGATACAAAG 207
51 GlySerLUARgPheARgAspLeuValProAsnTYrAsnProAspIleI 67
|||||
208 GGGTCGAGCGCTTCCGGACCTGCTACCACTAACCAACCCCGACATAAT 257
67 ePheLYAspGLuGlnAsnSerGlyAlaAspARgLeuMetThrGluArgC 84
|||||
258 CTTCAAAGATGAGGAGAAACAGCGCGCAGACCGCTCATGTACAGAGGTT 307
84 ySLySGluARgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGAGCGGGGTGAACGCTCTAGCCATCGGGGTGATGAACATGTGGCC 357
101 GlyValARgLeuARgValThrGluGlyTRpAspGluAspGlyHISAlaI 117
|||||
358 GGAAGTACGGCTGACTGATGAGAGGCTGGAGCAGACGCGCCACCAAGC 407
117 aGlnARgSerLeuHISTYrGluGlyARgAlaLeuAspIleThrThrSerA 134

|||||
408 ACAGATTAACCTCCACTACAGCAAGGCCGCTGCTGGACATACCAACGCTG 457
134 sPARgAspARgAsnLYSTyrGLyLeuLeuAlaARgLeuAlaValGluAla 150
|||||
458 ACCGTGACCGTATTAATATATGTTTGTGGCGCGCCTAGCTGTGGAAGCC 507
151 GLYPheSPTRPValTYrTYrGLuSerARgAsnHISValHISValSerVa 167
|||||
508 GGATTCACAGTGGGTCTACTACGAGTCCCGCACACCATTCACGATCGGT 557
167 lYsAlAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||
558 CAAAGCTGATTAACACTGCGCGGTCCGAGCGGAGGC 594

seq_name: gb_pat1:AR063081

seq_documentation_block:
LOCUS AR063081 1190 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5844079.
ACCESSION AR063081
VERSION AR063081.1 GI:5990772
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1190)
AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related thereto
JOURNAL Patent: US 5844079-A 2 01-DEC-1998;
FEATURES
source 1. 1190 Location/Qualifiers
BASE COUNT 194 a 371 c 399 g 226 t
ORIGIN

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:

US-08-900-220C-17_COPY_20_198 x AR063081 ..

Align seg 1/1 to: AR063081 from: 1 to: 1190

1 AlaGInSerCySGlyPProGlyARgGlyProValGlyARgARgTYrAl 17
|||||
58 GCCGAGAGCTGCGGGCGCGGCGGAGACCGGTGGCGCGCGCTTATGT 107
17 aArgLYGlnLeuValProLeuLeuTYrLYSGlnPheValProGlyValP 34
|||||
108 GCGCAACCACTTGCTGCTGCTATACAAAGCAAGTTGTGCCAGTATGC 157
34 roGluARgThrLeuGlyAlaSerGlyProAlaGluGlyARgValAlaArg 50
|||||
158 CCGAGCGGACCTTGCGCGAGTGGCCAGCGAGGGAGGATACAAAG 207
51 GlySerLUARgPheARgAspLeuValProAsnTYrAsnProAspIleI 67
|||||
208 GGGTCGAGCGCTTCCGGACCTGCTACCACTAACCAACCCCGACATAAT 257
67 ePheLYAspGLuGlnAsnSerGlyAlaAspARgLeuMetThrGluArgC 84
|||||
258 CTTCAAAGATGAGGAGAAACAGCGCGCAGACCGCTCATGTACAGAGGTT 307
84 ySLySGluARgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGAGCGGGGTGAACGCTCTAGCCATCGGGGTGATGAACATGTGGCC 357
101 GlyValARgLeuARgValThrGluGlyTRpAspGluAspGlyHISAlaI 117

```

|||||
358 GGAGTACGGCTAGTGTGACTGAGGCGTGGAGCGAGCGCGCACACGCC 407
117 aglnaspsrleuHISTYRGLUGLYARGLAalaaspIleThrSera 134
|||||
408 ACAGATTACTCTACACTACGAMAGCGCGCTTGACATCACCACTCTG 457
134 spArGspArGsnlySTYRGLyleuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGTACCGCTATAGTATGTTGTTGCGCGCCTAGCTGTGAAGCC 507
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSera 167
|||||
508 GGATTCGACTGGGTCTACTACGAGTCCCGAACACATCCACGTATCGGT 557
167 llysaAlaAspAsnSerLeuAlaValArgAlaGly 179
|||||
558 CAAAGCTGATTACTACTGCGGTCCGAGCGGAGGC 594
seq_name: gb_pat1:AX054856

```

```

seq_documentation_block: 1190 bp DNA PAT 13-JAN-2001
LOCUS AX054856
DEFINITION Sequence 3 from Patent W00073337.
ACCESSION AX054856
VERSION AX054856.1 GI:12228291
KEYWORDS
SOURCE
ORGANISM

```

```

human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1190).
AUTHORS
Peplinsky, R.B., Taylor, F. and Garber, E.
TITLE
Polymer conjugates of hedgehog proteins and uses
JOURNAL
Patent: WO 0073337-A 3 07-DEC-2000;
BIOGEN, INC. (US)

```

```

FEATURES
Location/Qualifiers
1..1190
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 194 a 371 c 399 g 226 t
ORIGIN

```

```

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

```

```

alignment_block:
US-08-900-220c-17_COPY_20_198 x AX054856 ..

```

```

Align seg 1/1 to: AX054856 from: 1 to: 1190

```

```

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgTyrAla 17
|||||
58 GCCCAGAGCTGCGCGCGCGCGAGCAGCGGTGCGCGCGCGGTATG 107
17 aArgGlyGlnLeuValProLeuLeuTyrGlnPheValProGlyValP 34
|||||
108 GCCCAAGCACTGTGCTCTGCTATACAGCAGTTGTGTCAGATATGC 157
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
158 CCGAGCGGAGCCCTGGCGCGAGTGGCCAGCGGAGGAGGAGTAAACAG 207
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle 67
|||||
208 GGCTCGAGCGCTTCCGCGACTCTGTAACCACTAACACCCCAATAT 257
67 ePheLysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArg 84
|||||
258 CTTCAGAGTGAAGAGAACAGCGCGCGAGCCCTGATGACAGAGCGTT 307

```

```

84 yslYsGluArgValaAlaAlaIleAlaValMetAsnMetTrpPro 100
|||||
308 GCAAGAGCGCGGTGAGACGCTCTAGCCATCGCGGTGATGAACATGTGAGCC 357
101 GlyValArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHisAla 117
|||||
358 GGAGTACCGCTAGCTGTGACTGAGGCTGGAGCGAGCGCGCACACGCC 407
117 aglnaspsrleuHISTYRGLUGLYARGLAalaaspIleThrSera 134
|||||
408 ACAGATTACTCTACACTACGAMAGCGCGCTTGACATCACCACTCTG 457
134 spArGspArGsnlySTYRGLyleuLeuAlaArgLeuAlaValGluAla 150
|||||
458 ACCGTACCGCTATAGTATGTTGTTGCGCGCCTAGCTGTGAAGCC 507
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSera 167
|||||
508 GGATTCGACTGGGTCTACTACGAGTCCCGAACACATCCACGTATCGGT 557
167 llysaAlaAspAsnSerLeuAlaValArgAlaGly 179
|||||
558 CAAAGCTGATTACTACTGCGGTCCGAGCGGAGGC 594
seq_name: gb_rol:MMDHH

```

```

seq_documentation_block: 1191 bp DNA ROD 18-JAN-1994
LOCUS MMDHH
DEFINITION M.musculus (129/Su) Dhh gene.
ACCESSION X76292
VERSION X76292.1 GI:443941
KEYWORDS
desert hedgehog protein; dhh gene.
SOURCE
house mouse.
ORGANISM
Mus musculus

```

```

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1191)
AUTHORS
McMahon, A.P.
TITLE
Direct Submission
JOURNAL
Submitted (24-NOV-1993) A.P. McMahon, Harvard University, 16
Divinity Ave., Cambridge, MA 02138, USA
2 (bases 1 to 1191)
AUTHORS
Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J.,
McMahon, J.A. and McMahon, A.P.
TITLE
Sonic hedgehog, a member of a family of putative signaling
molecules, is implicated in the regulation of CNS polarity
JOURNAL
Cell 75 (7), 1417-1430 (1993)
MEDLINE
94094334

```

```

FEATURES
Location/Qualifiers
1..1191
/organism="Mus musculus"

```

```

/strain="129/Sv"
/db_xref="taxon:10090"
/clone_lib="Stratagene; lambda fixIT"
1..24
/gene="Dhh"
1..1191
/gene="Dhh"
/codon_start=1
/product="desert hedgehog"
/protein_id="CAA53924.1"
/db_xref="GI:443942"
/db_xref="MGI:94891"
/db_xref="SPTREMBL:O61488"

```

```

/translacion="MALPASLPLDCCALLLALSAOSGCGPGKGRRRYKOLVPL
YKQVPSMPERTLGAAGPAERGVTRGSRFRDLVFNYPDILTFDEENSADRLWTER
CKERVNALAIAMNMPCVRLRYTEGMDSDHQAODSLHYGRALDITTSRDRNKYG
LLARLAVAGFDWYVESRNHIVSKADNSLAVAGSCFPGNATVRLRSRKGRLRE
LHRGDMVLADADAGVVPVPLVLDRLDRLRASVAVETERPPKLLTPWHIVFAA
RGPAPGDPAVPAVRRLRAGDSYVAPGCDLQPRVAVAREAVAGVAPATLACTL
LVNDYLASCYAVLVESHQWNAHRAFDRLRLHLAGLIPGAVOPTGMHYSRLTLPLAE
ELMG"

```


104 LeuArgValThrGluGlyTrpAspGluAspGlyHisHisAlaGlnAspSe 120
|||||
364 CTACGGGTTACGAGGGTTGGAGCAGATGACACACGACGTCACACAC 413
120 rLeuHisTyrGluGlyArgAlaLeuAspPheThrThrSerAspArgAspA 137
|||||
414 TCTGCATTTATGAAGGCGCGCACTGCACATTCACATCTGACCGAGACA 463
137 rGAsnTyrTrpGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPheAsp 153
|||||
464 GGAACAATAACGCGATGTGGCAGCTGCGACATGGAAGCTGCTTGAC 513
154 TrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAlaAs 170
|||||
514 TGGGTATATATGAATCCAAAGCACATCCATGCTCTGTGAAGACAGA 563
170 pAsnSerLeuAlaValArgAlaGlyGly 179
|||||
564 TAACCTCCCTGGGCTGCTCGCTGCTGCTG 591
seq_name: gb_ov:XLU26350
seq_documentation_block: 1197 bp mRNA VRT 05-JAN-1996
LOCUS XLU26350
DEFINITION Xenopus laevis hedgehog protein 4 (hh4) mRNA, complete cds.
ACCESSION U26350
VERSION U26350.1 GI:1147829
KEYWORDS
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodidae; Xenopus.
REFERENCE 1 (bases 1 to 1197)
Eker, S.C., McGrew, L.L., Lai, C.J., Lee, J.J., von Kessler, D.P.,
Moon, R.T. and Beachy, P.A.
Distinct expression and shared activities of members of the
hedgehog gene family of Xenopus laevis
Development 121 (8), 2337-2347 (1995)
JOURNAL 95401852
TITLE 2 (bases 1 to 1197)
REFERENCE 1 (bases 1 to 1197)
Eker, S.C., McGrew, L.L., Lai, C.L., Lee, J.J., von Kessler, D.P.,
Moon, R.T. and Beachy, P.A.
Direct Submission
TITLE Submitted (04-MAY-1995) Stephen C. Eker, Molecular Biology and
Genetics, Johns Hopkins University, 725 N. Wolfe St./714 PCIB,
Baltimore, MD 21205, USA
FEATURES
source
1. .1197
/organism="Xenopus laevis"
/db_xref="taxon:8355"
gene
1. .1197
/gene="hh4"
CDS
1. .1197
/gene="hh4"
/note="related to Xenopus laevis cephalic hedgehog,
encoded by GenBank Accession Number U26349; secreted
signaling molecule; Method: conceptual translation
supplied by author"
/codon_start=1
/product="hedgehog protein 4"
/protein_id="AA85164.1"
/db_xref="GI:1147830"
/translation="MPAVRLLIACCCMLLPVRCGPGRGPGVGGRRYRRLVPL
LYKQFVNPVPERKTGASGSGSKTRGSERTKLVNPNDITFRDENFTADRLMTE
RCRDVNAALISVNMMPGLKLVTEGWDGHHADSLHTEGRALDITSDRDNKY
GLARLVAEGFDWVYSESKAHIVSVNTDNGSLVRSQSGCPGTAMVMEGKKPLS
ELKLGDTFTDEGLIHLSVLLFHDHPYKATFVLEAGHPKTLVNPNLFTI
KSSSTGQPTFEAYRVOGLDIQVNTGTOSSKVVVSVDEQGYVAPMPEGTL
VDGYLTSCYATVESHTLHAHSLAPLRFLQGLASMLPDLHTSDGVNMYCHILYLAKY
LWMDMP"
BASE COUNT 295 a 286 c 314 g 302 t

ORIGIN
Alignment_scores:
Quality: 809.00 Length: 177
Ratio: 4.787 Gaps: 2
Percent Similarity: 95.480 Percent Identity: 84.746
Alignment_block:
US-08-900-220c-17_copy_20_198 x XLU26350 ..
Align seg 1/1 to: XLU26350 from: 1 to: 1197
4 CysGlyProGlyArgGlyProVal...GlyArgArgGlyArgGly 19
70 TGTGTCCTGGCGGGGGGGGGTGGCGGGCGGGCGGCTTACTATGAGAG 119
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGluA 36
120 A...CTGGTCCCACTGCTTTACAAAGCAGTTTGTGCTTAATGTGCCAGAGA 166
36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
167 AGACTCTAGGGGCGGAGTGCAGACTCAGAGGCGAGATCCGACGGGCTCG 216
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLe 69
217 GAAAGTTCATCAAGTTGGTCCCACTAACCCAGACATCATCTTCA 266
69 sAspGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
267 GGATGTGAAGAACACACAGAGCTGACCTCATCATGATGAGCGCTGCAGAG 316
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
317 ACAGACTCATAGCTTGGCTGCTGTGTGTGTAACATGTGGCCAGGCTTG 366
103 ArgLeuArgValThrGluGlyTrpAspGluAspGlyHisHisAlaGlnAs 119
367 AACCTCGGGGTGACGAGGGCTGGATGAGAGGCGCACACGCTCAGCA 416
119 pSerLeuHisTyrGluGlyArgAlaLeuAspPheThrThrSerAspArgA 136
417 CTTCTTGATTTATGAAGGCGGACCTGTGCACATTAACATCTGACCGAG 466
136 sPrpArgAsnTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
467 ACAGGAACAATATGATGATGTCGCCGCTGACAGTGAAGCTGCTGTT 516
153 AspTrpValTyrTrpGluSerArgAsnHisValHisValSerValLysAl 169
517 GACTGGGTGATTAACGATCCAAAGCACACATCATGCTCTGTGAACAC 566
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
567 AGATAACTCCCTCGGCTGCTCGCTGCTGCTG 597
seq_name: gb_r01:MMU85610
seq_documentation_block: 2103 bp mRNA ROD 16-MAR-1997
LOCUS MMU85610
DEFINITION Mus musculus Indian hedgehog protein (Ihh) mRNA, complete cds.
ACCESSION U85610
VERSION U85610.1 GI:1890096
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2103)
Valentin, R.P., Brookhiser, W.T., Park, J.J., Yang, T., Briggs, J.,
Dressler, G. and Holman, L.B.
Post-translational Processing and Renal Expression of Mouse Indian

Hedgehog
J. Biol. Chem. (1997) In press
2 (bases 1 to 2103)
Valentin, R.P., Brookhiser, W.T., Park, J., Yang, T., Briggs, J.,
Dressler, G. and Holzman, L.B.
TITLE
Direct Submission
JOURNAL
Submitted (15-JAN-1997) Internal Medicine/Division of Nephrology,
University of Michigan Medical School, 1560 MSRB II--Box 0676, Ann
Arbor, MI 48109-0676, USA
FEATURES
Location/Qualifiers
1. 2103
/organism="Mus musculus"
/strain="CD-1"
/db_xref="taxon:10090"
184..1533
/gene="Inh"
184..1533
/gene="Inh"
/codon_start=1
/product="Indian hedgehog protein"
/protein_id="AA049692.1"
/db_xref="GI:1890097"
/translation="MESPRAQTQPSPKLSOPRAHLAHOAPSPALPGYPAMPAMLT
RRRLRCLLILLIPARAGCGPRVYSGRRPRKLYPLAKYKQSPVVPKTKGAS
GREESKIASSESRFKELTPNYPDIIFDEENTGADRLTQCKDRLNSLAISNMOM
PEKLVTEGWDESDHSESLHYEGRAVDITSDRNKYGILRLAIVEAGFDWVY
ESKAVHGSVSKSEHSAAKTGGCFPAGAVLENERVALSAVKPGRVLAEGEDGP
TFSVDVLEDRPNRLRAFOVETODPPRLALTPAHLEIFIDNHTPEPAHRAFEAS
HYQPGQYVLSGVPGLQPARVAASVTHVALGSAVAPLRIHGLVEEDVVASCFRAVADH
HLAOLAFMPLRFPISLAWSMTPSEGVHYPQMIYRLGLLLESTFHLGMSGAS"

BASE COUNT 376 a 684 c 628 g 415 t
ORIGIN

alignment_scores:
Quality: 743.50 Length: 179
Ratio: 4.479 Gaps: 1
Percent Similarity: 92.737 Percent Identity: 75.419

alignment_block:
US-08-900-220C-17_COPY_20_198 x MM085610 ..
Align seg 1/1 to: MM085610 from: 1 to: 2103

```
1  Alaglnsercysgilyproglyargilyprovalajlyargargatyrval 17
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
370 GCGCGGGGCTGCGGGCGGGCGGGCGG...GTGCTGGCAGCGCCCGGAGGCC 416
17 aarglyglnleuvalproleuleutyrlysglnphevalproglyvalp 34
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
417 GCCTCGCAAGCTCGTCCCTTCTGCTTACAGACATTCAGCCCAAGCTGC 466
34 rogluatgrhleuglyalasercilyproalaglucilyargvalalaarg 50
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
467 CGGAGAGACCCCTGGGGCGGCGGCGGCTACGACGACGACAGATGCCGCG 516
51 glysersilyargpheargaspheulvalproasntyrasnproasplleil 67
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
517 AGCTCTGAGGCTTCAAGAGCTCACCCCAACATCAATCCCGACATCAT 566
67 ephelyaspglulnansercilyalaaspargleumetthrgluaragc 84
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
567 CTTCAAGAGACGAGAGACACGGGTGCGACCGCTCATGACCCGAGGCTT 616
84 yslsglyuarvalasnalaleualalalealavalmetasmettrp 100
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
617 GCAAGACCGCTCTACACTGACCTCTGTGATGATGACCAACCATGCGCT 666
101 glyvalargleuargvalthrgluglytrpaspilnaspilyhshsai 117
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
667 GGTTGAAGTGGCGGTGACCGAAGGCTGGGATGAAGATGCGCATCTACTC 716
117 aglinspserleuhtsyrngluglyargalaleuaspllethrthrsra 134
```

```
717 AGAGAGCTCTTACACTATGAGGGCGCGGTGATATACACCACTCAG 766
134 spagaspargasnlystyrnglyleuvalalaargleualalaigla 150
767 ACCGTGACCGAAATAGATAGTACTGCTGGCGGCGCTTACGATGGAGCC 816
151 glypheasprtpvaltyrtyrglusertargasnhsvalhshsval 167
817 GGCTTCGACTGGGTGTATACGAGTCCAGGCCACGCTGATCTCTGT 866
167 llyalaspasnsrleuvalalealargalagly 179
867 CAAGCTGAGCATTCGGCGCTGCCAAGACAGGTGC 903
seq_name: gb_pat1:AR021198
seq_documentation_block: 1277 bp DNA PAT 05-DEC-1998
LOCUS AR021198
DEFINITION Sequence 1 from patent US 5789543.
ACCESSION AR021198
VERSION AR021198.1 GI:3975813
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Ingham, P.W., McMahon, A.P. and Tablin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins and uses related
JOURNAL Patent: US 5789543-A 1 04-AUG-1998;
FEATURES Location/Qualifiers
1..1277
source
/organism="unknown"
BASE COUNT 262 a 396 c 387 g 232 t
ORIGIN

alignment_scores:  
Quality: 741.00 Length: 177  
Ratio: 4.411 Gaps: 2  
Percent Similarity: 94.915 Percent Identity: 74.011



alignment_block:  
US-08-900-220C-17_COPY_20_198 x AR021198 ..  
Align seg 1/1 to: AR021198 from: 1 to: 1277



```
3 sercysgilyproglyargilyprovalajlyargargatyrvalaargly 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
76 ACTGTGGACCGAGGAGGGC...ATTGGAAAAAGAGGCGACCCCAAAA 122
19 sglleuvalproleuleutyrlysglnphevalproglyvalproglua 36
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
123 G...CTGACCCCTTACCTTAAAGCACTTATTCCTCAATGTGCACAGNA 169
36 rgrthleuglyalasercilyproalaglucilyargvalalaarglyser 52
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
170 AGACCTTAGGGGCGCAGTGAAGATATGAAGGAGACATCAACAAAGACTCC 219
53 gluatrgpheargaspheulvalproasntyrasnproaspllelephely 69
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
220 GAGGATTTTAAAGAACTAACCCCAATTTACAACTTATATTTTAA 269
69 saspglulnansercilyalaaspargleumetthrgluaragcylasg 86
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
270 GGATGAAGACACACCGGAGCTGCACACATGATGATCAGGCTGCAAGG 319
86 luargvalasnalaleualalalealavalmetasmettrp 102
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
320 ACAAGCTGAATGCCCTGCGATCTCGGTGATGAACACAGTGGCGGGGTG 369
103 argleuargvalthrgluglytrpaspilnaspilyhshsaiaglnas 119
```


```

370 AAGCTGGGGGACCGAGGCGCTGGAGCAGAGATGGCCATCTCCGAGGA 419
119 pserleuhtstygltgluylarvalaleuaspilletthrsersaparga 136
420 ATCGGTGACACAGAGGGTCCGGCCGTGCATCCACAGCTGGATCCGG 469
136 spargAsnlystygltgluylarvalaleuaspilletthrsersaparga 152
470 ACCGAGCAGAGTACGAGATGCTGGCCGCTGCGTCGAGCGCGCTTC 519
153 AsptPvalTYTyrTgluSerArgSnhSvalHsValSerValLysAl 169
520 GACTGGGTCTACTACGAGTCCAGAGCGCACATCCACTGCTCCGTAAGC 569
169 aaspasnsrleuAlaValArgAlaGlyGly 179
570 AGAAACTCAGTGGCCAGCAATCAGAGAGC 600

seq_name: gb_pat1:AR063080

seq_documentation_block:
LOCUS AR063080 1277 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5844079.
ACCESSION AR063080
VERSION AR063080.1 GI:5990771
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Ingham, P.W., McMahon, A.P. and Tablin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related
thereo
JOURNAL Patent: US 5844079-A 1 01-DEC-1998;
FEATURES
source Location/Qualifiers
1..1277
BASE COUNT 262 a 396 c 387 g 232 t
ORIGIN

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.411 Gaps: 2
Percent Similarity: 94.915 Percent Identity: 74.011

alignment_block:

US-08-900-220c-17_COPY_20_198 x AR063080 ..

Align seg 1/1 to: AR063080 from: 1 to: 1277

3 SerCysGlyProGlyArgGlyProValIglyArgArgArgTyrAlaArgly 19
76 ACTTGTCGACGACGAGGGGC...ATTGGAAAAAGAGGACCCCAAAAA 122
19 scInleuValProLeuLeuTyrLysGlnPheValProGlyValProGlyua 36
123 G...CTGACCCCGTTAGCCTATAGCAGATTATTCCTCATGTGGCGAGAA 169
36 rgtHrleuGlYAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
170 AGACCTAGGGGCGCAGTGAAGATATGAAGGAGATCACAAGAAACTCC 219
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
220 GAGAGATTAAAGAACTAACCCCAATTACACCCGACATTATTTTAA 269
69 sAspGluGlnAsnSerGlyAlaAspArgLeuMetHrGluArgCysLysG 86
270 GGATGAGAGAAACGCGGAGCTGACAGACTGATGACTACGCGCTGAAAG 319
86 luArgValAsnAlaLeuAlaIleAlaValaMetAsnMetTyrProGlyVal 102

320 ACAACCTGAATGCCCTGGGATCTCGGTATGAACACAGTGGCGGGGTG 369
103 ArgLeuArgValThrGluGlyTyrAspGluAspGlyHsHsAlaGlnAs 119
370 AAGCTGGGGGTACCGAGGGCTGGAGCAGAGATGGCCATCTCCGAGGA 419
119 pserleuhtstygltgluylarvalaleuaspilletthrsersaparga 136
420 ATCGGTGACACAGAGGGTCCGGCCGTGCATCCACAGCTGGATCCGG 469
136 spargAsnlystygltgluylarvalaleuaspilletthrsersaparga 152
470 ACCGAGCAGAGTACGAGATGCTGGCCGCTGCGTCGAGCGCGCTTC 519
153 AsptPvalTYTyrTgluSerArgSnhSvalHsValSerValLysAl 169
520 GACTGGGTCTACTACGAGTCCAGAGCGCACATCCACTGCTCCGTAAGC 569
169 aaspasnsrleuAlaValArgAlaGlyGly 179
570 AGAAACTCAGTGGCCAGCAATCAGAGAGC 600

seq_name: gb_pat1:AX054854

seq_documentation_block:
LOCUS AX054854 1277 bp DNA PAT 13-JAN-2001
DEFINITION Sequence 1 from Patent WO0073337.
ACCESSION AX054854
VERSION AX054854.1 GI:12228290
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Pepinsky, R.B., Taylor, F. and Garber, E.
TITLE Polymer conjugates of hedgehog proteins and uses
JOURNAL Patent: WO 0073337-A 1 07-DEC-2000;
FEATURES
source Location/Qualifiers
1..1277
BASE COUNT 262 a 396 c 387 g 232 t
ORIGIN

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.411 Gaps: 2
Percent Similarity: 94.915 Percent Identity: 74.011

alignment_block:

US-08-900-220c-17_COPY_20_198 x AX054854 ..

Align seg 1/1 to: AX054854 from: 1 to: 1277

3 SerCysGlyProGlyArgGlyProValIglyArgArgArgTyrAlaArgly 19
76 ACTTGTCGACGACGAGGGGC...ATTGGAAAAAGAGGACCCCAAAAA 122
19 scInleuValProLeuLeuTyrLysGlnPheValProGlyValProGlyua 36
123 G...CTGACCCCGTTAGCCTATAGCAGATTATTCCTCATGTGGCGAGAA 169
36 rgtHrleuGlYAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
170 AGACCTAGGGGCGCAGTGAAGATATGAAGGAGATCACAAGAAACTCC 219
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
220 GAGAGATTAAAGAACTAACCCCAATTACACCCGACATTATTTTAA 269

```

69 saspGlucIuansSerGlyAlaaspArgLeuMetThrGluArgCysIysG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 GGATGAAGAGAACACGGAGCTGACAGACTGATGACTACAGCGCTGCAAG 319
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 AACAGCTGAATGCCCTGGCATCTGGTGGATGAACACAGTGGCCGGGGTG 369
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AACCTGGGGGAGACCGAGCGCTGGAGAGATGGCCATCCTCCGAGGA 419
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATCCCTCAGCTACGAGGGTGGCGCGCTGGACATCCACACGCTCGGATCGG 469
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCCGACCAAGTACGGAATGCTGGCCCGCTCGCGTGGAGCGCGCTTC 519
153 AspTyrValTyrTyrGluSerArgAsnHisValHisValSerValIysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GACTGGGTCTACTACGAGTCCCAAGGGCCACATCCACTGCTCCGCAAGC 569
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 AGAAACTCAGTGGCAGCAATCAGAGGCG 600

seq_name: gb_ov:CHKZPAMED

seq_documentation_block:
LOCUS CHKZPAMED 1567 bp mRNA VRT 04-OCT-1999
DEFINITION Gallus gallus sonic hedgehog mRNA, complete cds.
ACCESSION L28099
VERSION L28099.1 GI:453526
KEYWORDS
SOURCE
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1567)
REFERENCE
Riddle, R.P., Johnson, R.L., Laufer, E. and Tabin, C.
Sonic hedgehog mediates the polarizing activity of the ZPA
JOURNAL
Cell 75, 1401-1416 (1993)
94094333
FEATURES
source
location/Qualifiers
1..1567
/organism="Gallus gallus"
/db_xref="taxon:9031"
/issue_type="embryo"
/dev_stage="stage 20-22"
227..1504
/codon_start=1
/product="sonic hedgehog"
/protein_id="AA072428.1"
/db_xref="GI:453527"
/translation="MVEMLLRLRLVGFICALLVSGLTGPGRGIGKRHPKRLTP
LAYKOFIYNVAEKTIGASRGREKLTIRNSERKELTPVNPDIIRKDEENQDRLMT
ORCKDKLAIALISVNMOMGVKLRVTEGMDHSGESLSHREGRAVDITNSDRLSK
YGMRLARLAVEGFDPVYTESKAHILCSYKAENSVAKSGCCPGSATYHLEHGTRKL
KDLSPGVDNLADADGRLLYSDLEFLDRLMDSRLFLYIETRQRRARLLTAHLLF
VAPHNOSFATSGOALFASNVKPGQGVVYLGSGOOLPASVHVSLLREASGAY
APLTAQGTILINRVASCAYVIEESHMAHMAFAPRLAOGILALCPDCAIPTAATTT
TGIIHYSRLRYHIGSWILDGALHPLGMVAPAS"
BASE COUNT 347 a 457 c 475 g 288 t
ORIGIN

```

```

alignment_scores:
Quality: 741.00 Length: 177
Ratio: 4.411 Gaps: 2

```

```

Percent Similarity: 94.915 Percent Identity: 74.011
alignment_block:
US-08-900-220C-17_COPY_20_198 x CHKZPAMED
Align seg 1/1 to: CHKZPAMED from: 1 to: 1567

```

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaArgGly 19
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 ACTTGTGGACGACGACGAGGCG...ATTGAAAAAGGAGGACCCCAAAA 348
19 sGlnLeuValProLeuLeuTyrIysGlnPheValProGlyValProGluA 36
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
349 G...CTGACCCCGTTAGCCCTATAGCAGTTATTTCCAAATGTGGCAGAGA 395
36 rGThrLeuGlyAlaSerGlyProAlaGluArgValAlaArgGlySer 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
396 AGACCCCTAGGGCCGACGTGAAGATATGAAGGAAGATCACAACTCC 445
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
446 GAGAGATTATTAAGAACTAACCCCAATTACACCTGACATTTATTTTAA 495
69 saspGlucIuansSerGlyAlaaspArgLeuMetThrGluArgCysIysG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 GGATGAAGAGAACACGGAGCTGACAGACTGATGACTACAGCGCTGCAAG 545
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
546 ACAAGCTGAATGCCCTGGCATCTCGGTGATGAACACAGTGGCCGGGGTG 595
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
596 AAGCTGGGGTGGACCGAGGCTGGAGAGATGGCCATCCTCCGAGGA 645
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
646 ATCGCTCAGCTACGAGGCTGGCGCGCTGGACATCACACGTCGGATCGG 695
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
696 ACCGACCAAGTACGGAATGCTGGCCCGCTCGCGTGGAGCGCGCTTC 745
153 AspTyrValTyrTyrGluSerArgAsnHisValHisValSerValIysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
746 GACTGGGTCTACTACGAGTCCCAAGGGCCACATCCACTGCTCCGCAAGC 795
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
796 AGAAACTCAGTGGCAGCAATCAGAGGCG 826

seq_name: gb_pat1:AX054866

seq_documentation_block:
LOCUS AX054866 1622 bp DNA PAT 13-JAN-2001
DEFINITION Sequence 13 from Patent WO0073337.
ACCESSION AX054866
VERSION AX054866.1 GI:12228296
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1622)
REFERENCE
Pepinsky, R.B., Taylor, F. and Garber, E.
Polymer conjugates of hedgehog proteins and uses
JOURNAL
Patent: WO 0073337-A 13 07-DEC-2000;
BIOGEN, INC. (US)
FEATURES
source
location/Qualifiers
1..1622
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 277 a 549 c 510 g 286 t

```

ORIGIN

[illegible]

alignment_block:

US-08-900-220C-17_COPY_20_198 x AX054866

Align seg 1/1 to: AX054866 from: 1 to: 1622

1 ALaGInSerCySLyPProGlyAArgGlyProValGlyIArgArgTyrAla 17
123 GCATGGGGCTGGCGGGCCGGCTCG...GTGGTGGGACAGCCGGCGGACG 169
17 aArgLySGInLeuValProLeuLeuTyrLySGInPheValProGlyValP 34
170 GCCAGGCAAACTCGTGGCTGGCTGCCTCAAGACATTCAGCCCAATGTGG 219
34 roGluArgThrLeuGlyValaSerGlyProaGluGlyArgValAlaArg 50
220 CCGAADAACCTGGGCGCCAGCGGAGCGTATGAGGAAGCAAGTCGTCGG 269
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
270 AGCTCCGAGCCGCTTCAAGAGACTACCCCAATTACATCCACACATCAT 319
67 ePheLyAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
320 CTTCAAGAGACGAGAGAACACAGCGCGCCAGCCGCTCATGACCCAGCGCT 369
84 yLySGluArgValaAsnAlaLeuAlaIleAlaValaMetAsnMetTPro 100
370 GCAAGAGCCGCTGAACTCGTGGCTATCTCGGTGATGAACCAAGTGGCC 419
101 GlyValaArgLeuArgValaThrGluGlyTyrAspGluAspGlyIHisAla 117
420 GGTGTGAAGCTCGCGGTACCCGAGCGGTGGAGACAGAGCCCAACCATTC 469
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSera 134
470 AGAGAGATCCCTGCATTATGAGGGCGCGCGGTGGACATCACACATCAG 519
134 sPArgAspArgAsnLySTyrGlyLeuLeuAlaArgLeuAlaValaGluAla 150
520 ACCGCGAGCCGCAATTAAGTATGACTGTGGCGGCTGGAGTGGAGGCGC 569
151 GlyPheAspTyrValaTyrTyrGlySerArgAsnHisValaHisValaSerA 167
570 GGCTTTGACTGGGTGTATTAAGACATCAAAAGGCCCAAGTCATGTCTCCGT 619
167 lLyAlaAspAsnSerLeuAlaValaArgAlaGlyGly 179
620 CAAATGCCAGACACTCGGCGGAGCCAGACAGCGGCGGC 656
seq_name: gb_pat1:AR021201

seq_documentation_block:

LOCUS	AR021201	1313 bp	DNA	PAT	05-DEC-1998
DEFINITION	Sequence 7 from patent US 5789543.				
ACCESSION	AR021201				
VERSION	AR021201.1 GI:3975816				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 1313)				
AUTHORS	Ingham,P.W., McMahon,A.P. and Tabin,C.J.				
TITLE	Vertebrate embryonic pattern-inducing proteins and uses related thereto				
JOURNAL	Patent: US 5789543-A 7 04-AUG-1998;				

FEATURES

FEATURES	Location/Qualifiers
source	1. 13j3
BASE COUNT	/organism="unknown"
ORIGIN	254 a 411 c 428 g 220 t

```
alignment_scores:
  Quality: 740.00
  Ratio: 4.431
  Percent Similarity: 94.350
  Length: 1777
  Gaps: 2
  Percent Identity: 74.576
```

alignment_block:

US-08-900-220C-17_COPY_20_198 x AR021201

Align seg 1/1 to: AR021201 from: 1 to: 1313

```

19 scInLeuValProLeuLeuTyrLysGlnPheValProGlyValProGluA 36
17 G...CTGACCCCTTTAGCCTACCAAGAGATTATTTCCCAACGTAGCCGAGA 163
16 rGThrLeuGlyAlaSerGlyProAlaGlnGlyAlaGlyAlaArgGlySer 52
154 AGACCTTAGGGGCCAGCGGAGATATGAAAGGAAAGATCAACAAGAAATCC 213
53 GluArgPheArgAspLeuValProAsnTyrAspProAspIleIlePhe 69
214 GAGAGATTTAGGAACATCACCCCAATTACAAACCCGACATCATATTTA 263
69 SASPGluGluAsnSerLysAlaAspArgLeuMetThrGluArgCysLys 86
264 GGATGAGGAAACACGGGACACAGCCGCTGATGACTCGAGGTCCAAG 313
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTyrProGlyAl 102
314 ACAAGTTAAATGCCCTTGCCCATCTCTGTATGAAACCAAGTGGCTGAGTG 363
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
364 AGGCTGGCAGTAGCCGAGGGCTGGGATGAGGACGGCCATCATTCAGAGGA 413
119 PserLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgA 136
414 GTCTCTACACTATGAGGGTGCAGACAGTGCACATCCACAGTCCGACCGGG 463
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGlyPhe 152
464 ACCCGACGAAAGTACGGCATCGCTGGCTGGCTGGTGGAAACAGAGTTTC 513
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
514 GACTGGGTCTACTATGATCCAAAGCTCAATCCACTGTTCTGTGAAGC 563
169 aAspAsnSerLeuAlaValAlaArgAlaGlyLys 179
564 AGAGAACTCCGTGGCGGCCCAATTCGGCGGC 594

seq_name: gb_pat1:AR063083

seq_documentation_block:
LOCUS AR063083 1313 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 4 from patent US 5844079.
ACCESSION AR063083
VERSION AR063083.1 GI:5990774
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1313)

```

AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related thereto

JOURNAL Patent: US 5844079-A 4 01-DEC-1998;
FEATURES Location/Qualifiers
source 1. 1313

BASE COUNT 254 a 411 c 428 g 220 t
ORIGIN

alignment_scores:
Quality: 740.00 Length: 177
Ratio: 4.431 Gaps: 2
Percent Similarity: 94.350 Percent Identity: 74.576

alignment_block:
US-08-900-220c-17_copy_20_198 x AR063083 ..

Align seg 1/1 to: AR063083 from: 1 to: 1313

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAlaArgly 19
  ::::::::::::::::::::
70 GCGTGTGGGCCCGCCAGGGG...TTGGAAAGAGCGCGCACCCCAAAA 116
  ::::::::::::::::::::
19 SGInLeuValProLeuLeuTyrTrpGlnPheValProGlyValProGlu 36
  ::::::::::::::::::::
117 G...CTGACCCCTTACCTACAGCAGTTATTCCCAAGTACCGAGA 163
  ::::::::::::::::::::
36 rGTrhleuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
  ::::::::::::::::::::
164 AGACCCCTAGGGGCCAGCGCAGATATGAGGAGGAGATCAGAGAACTCC 213
  ::::::::::::::::::::
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
  ::::::::::::::::::::
214 GAACGATTTAAGGAACCTACCCCAATTACACCCGACATCATATTAA 263
  ::::::::::::::::::::
69 SAARGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  ::::::::::::::::::::
264 GGATGAGGAAACACGCGGACAGCCGCTGATGACTCAGAGGTGCAAG 313
  ::::::::::::::::::::
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  ::::::::::::::::::::
314 ACAAGTTAAATGCCCTTGCCATCTCTGATGAACCACTGCCCTGAGTG 363
  ::::::::::::::::::::
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAlaGlnAs 119
  ::::::::::::::::::::
364 AGGCTGCGAGTGAACCGAGCGCTGGATGAGAGCGGCATTCACAGAGA 413
  ::::::::::::::::::::
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
  ::::::::::::::::::::
414 GTCTTACACTATGAGGCTGAGCAGTGCATCACACGTCGCCACCGGG 463
  ::::::::::::::::::::
136 SPARGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
  ::::::::::::::::::::
464 ACCGCAACAGTACGCGATGCTGGCTGCTGGCTGGAACGAGTTTC 513
  ::::::::::::::::::::
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
  ::::::::::::::::::::
514 GACTGGGTCTACTATGATCCAAAGTCACATCCACTGTCTGTGAAGC 563
  ::::::::::::::::::::
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
  ::::::::::::::::::::
564 AGAGAACTCGTGGCGGCNAATCCGGCGG 594
  ::::::::::::::::::::

```